EVALUATION IN HEALTH EDUCATION AND PHYSICAL EDUCATION

McGraw-Hill Series in Health Education,
Physical Education, and Recreation

CLIFFORD LEE BROWNELL, *Consulting Editor*

ANDERSON · Industrial Recreation

BROWNELL · Principles of Health Education Applied

EVANS, BACON, BACON, AND STAPLETON · Physical Education for Elementary Schools

KENNEY AND LAW · Wrestling

KNAPP AND HAGMAN · Teaching Methods for Physical Education

KRAUS · Recreation Leader's Handbook

LAWRENCE AND FOX · Basketball for Girls and Women

METHENY · Body Dynamics

MILLER · Fundamentals of Track and Field Coaching

PATERSON · Foundations of Community Health Education

TORNEY · Swimming

WILLGOOSE · Evaluation in Health Education and Physical Education

EVALUATION IN HEALTH EDUCATION AND PHYSICAL EDUCATION

Carl E. Willgoose, Ed.D.

SCHOOL OF EDUCATION
BOSTON UNIVERSITY

McGRAW-HILL BOOK COMPANY, INC.
New York Toronto London
1919

EVALUATION IN HEALTH EDUCATION AND PHYSICAL EDUCATION

 Library of Congress Catalog Card Number 60-53358

THE MAPLE PRESS COMPANY, YORK, PA.

70310

Editor's Introduction

Over the years appropriate *evaluation* has remained as one of the continuous and perplexing problems that face conscientious educators. Authorities in various fields have wrestled with this problem, including numerous leaders in health education and physical education

Sound evaluation takes into account the objectives sought and the media through which they are achieved First, every educational institution has over-all objectives it strives to fulfill Second, each approved subject or program in that institution has its own definitive objectives that contribute toward the fulfillment of complete education through the media employed by a given subject or program For example, mathematics uses such media as numbers, symbols, and formulas, physical education embraces games and sports, aquatics, self-testing events, and other activities, and health education deals with functions and needs of the human organism that pertain to optimum and wholesome living for the individual, his family, and his community

But differences of opinion sometimes arise between the exponents of objective versus subjective evaluation Those who lean toward objective measures strive for a numerical score resulting from tests that meet the accepted criteria of objectivity, validity, and reliability Persons favoring subjective evaluation contend that many educational objectives of great worth do not lend themselves to statistical treatment Thus running speed or arm strength may be measured by objective tests of high validity and reliability; whereas attitudes toward play and recreation or toward habits of healthful living can be best appraised by subjective evaluation performed by competent observers Intelligent use of both objective and subjective evaluation, applied in their rightful place, seems best suited for the complete task at hand

Dr. Willgoose brings to health education and physical education a thorough understanding of the many facets of evaluation and experience in dealing with the variety of techniques associated therewith. This excellent text prepared by him should provide valuable assistance to prospective and in-service teachers who desire practical knowledge about the subject and who recognize the true worth of evaluation as a significant component of total education.

Clifford Lee Brownell,
CONSULTING EDITOR

Preface

Almost a half century ago some persons doubted that educational outcomes could be evaluated. It was this attitude that prompted Edward Lee Thorndike to make his famous statement that whatever exists at all exists in some amount and can be measured

Man as a physical and biological scientist has demonstrated for centuries his acceptance of Thorndike's point of view Man as an educational scientist, however, has been slow to apply this reasoning to educational outcomes Educational issues were settled and principles and policies were arrived at by rhetoric and reputation Research, involving appropriate measurement and appraisal, came along later

Despite the large amount of test materials and measurement devices available today in health education and physical education, there is a noticeable lack of measurement and evaluation practices in many schools throughout the country Much of this may be because a large number of health and physical educators are not convinced of the need for proper evaluation, despite undergraduate and graduate college courses in this area. There remains, therefore, the challenge to make these college courses so meaningful that the graduate will be enthusiastic about measuring his progress toward educational goals

This book attempts to present the more common areas of measurement and evaluation in such a manner that the professional student, with the help of reference books and other selected readings, will understand the value of measurement and evaluation by personally thinking and acting on problematical situations An effort has been made to afford the student a chance to think about and work with tests with a view toward their practical application

In general it is the aim of the book (1) to assist the future teacher in understanding the place and importance of measurement in the teaching process; (2) to provide a way for students to examine more closely some of the tools and other elements of measurement and evaluation, so they may become equipped to evaluate scientifically the school program in terms of educational aims and objectives; (3) to embrace in a practical manner the fundamentals of appraisal and to relate these to curriculum content, method of instruction, and pupil advancement; (4) to provide the student

with sufficient understanding to realize that the more valid the measurement and the more judicious the evaluation of results, the sooner the goal may be attained, and (5) to show how measurement may be used for the determination of pupil status, the classification of students, individual analysis and guidance, appraisal of teacher and program efficiency, and gaining the support of the administration and the community

A definite plan of construction was followed in building this text It was intended that the work should develop logically from the first to last chapters, that is, from the more fundamental considerations to the more detailed ones The first three chapters are introductory in nature Objectives are considered in Chapter 2 Test criteria and some test evaluation are considered in the following two chapters

Chapters 5 through 15 deal primarily with knowledge, habits, attitudes, and skills Chapters 4 and 5 essentially concern health education Chapters 6 through 14 are more relevant to physical education than to health education. There is considerable overlapping in the areas of health and physical education; thus certain chapters apply to both areas Certainly the chapters in which social efficiency, individual differences, administration, and research are considered are not limited to one area or the other but concern persons in both areas

Chapter 17 is presented to help the pupil appreciate the unique position of the administrator in the evaluation process

Throughout the book an attempt is made to make the student conscious of health deviations He is asked to observe behavior relative to the eyes, ears, general appearance, etc Such items are a means by which he is able to "detect and refer" The more signs of health abnormalities a teacher is familiar with, the more readily pupils in need of special attention may be brought to the school nurse, psychologist, or physician

The program that health and physical education personnel will create in the years to come will depend upon a dedicated leadership—a leadership thoroughly grounded in the philosophy and goals of health and physical education and the means for attaining them Moreover, this leadership will need periodically to appraise its progress and adapt programs to the needs of individual pupils Preparation in measurement and evaluation at both undergraduate and graduate levels is involved To this end the book is dedicated

The author wishes gratefully to acknowledge the numerous persons involved in research and evaluation whose ideas, tests, and materials have been employed in the construction of this book Thanks are extended to Mrs. Helen J. Gilbert for her long hours typing the manscript and to my wonderful family for their patience.

Carl E. Willgoose

Contents

CHAPTER 1

The Nature of Measurement and Evaluation

Man is the measure of all things.
Protagoras

From his first gasp of breath to his last expiration man works, struggles, and strives for the utopia of tomorrow. Directly and indirectly he *measures* his failures and successes To do this he uses complex measuring devices, such as the electronic computer in the business world, the seismograph in earth science, and the scintillation counter for recording impulses made by fast-disintegrating radioactive isotopes Without these and other measuring devices our present society would be reduced to the culture of the Stone Age.

There is order in the universe Man depends on it He is continually trying to adapt himself to it by "measuring up" to many things When he feels that he has regulated himself properly, he is no longer disturbed and he feels secure. To arrive at this desired state of adjustment, he may try to appraise himself, thinking in terms of his life goals He asks himself what he wants out of living What things are important? What knowledge is of most worth? Where does he want to go? Is he headed in the right direction? With such an approach to living, everyday activities take on greater meaning Man becomes more careful; he gauges his progress and regulates his behavior

Meaning of Measurement and Evaluation

People who have definite objectives and a desire to progress frequently feel frustrated from time to time A lack of knowledge of personal status may cause frustration This feeling of frustration is the primary reason for measurement. *Measurement, therefore, can be a process of making comparisons and relating them to personal needs in an effort to find out where one is headed.* Information, whether pleasant or otherwise, helps relieve frustration.

To evaluate, says Webster, is "to appraise carefully," and to appraise is "to set a value on." Value is determined by relative "worth, excellence, or importance." *The process of evaluating requires judgment of sufficient dis-*

criminatory power to compare facts, to perceive their relationships and attributes, and thus to distinguish truth from falsehood.

Knowledge of one's progress requires some degree of discriminative ability. When all the factors associated with human progress are carefully weighed and when man differentiates between them to ascertain which are the more important, a certain degree of discriminative ability has been demonstrated. It is the kind of ability that teachers and administrators show when they judge that educational advancement has taken place. Yet progress toward the goals of education is not inevitable; the educator must periodically take note of his status and that of his program and pupils. For example, in health and physical education, a good teacher does not assume that a healthy or well-adjusted school child is a natural outcome of exposure to classes in health and physical education. He knows that is the outcome only when certain learning conditions exist. These include, among

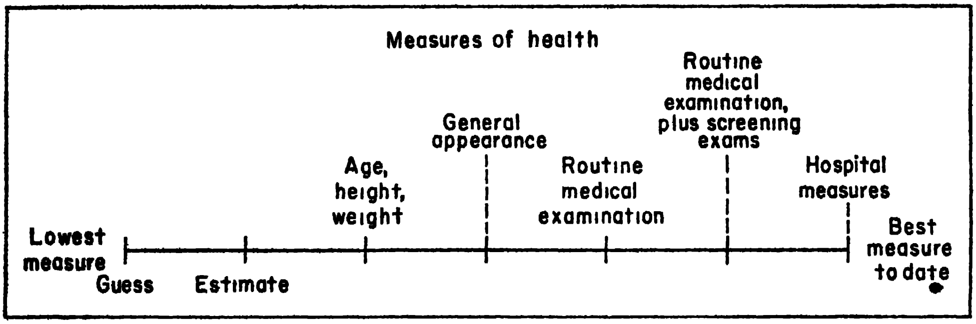


Fig. 1-1. Degrees of Health Measures

others, a curriculum fashioned in the light of preconceived goals, the methods and materials used, and the efficiency of the school staff.[1]

The lowest form of measurement is a pure guess Measurement is made more accurate by an estimate and improved considerably by instruments. For example, what is the length of a blackboard? By pure guess, it might be 5 feet. By estimate, in which some judgment or comparison is involved, it might be 7 feet. By measurement with a steel tape, it might actually prove to be 7 feet 6 inches in length. The same reasoning can be applied to a health measurement.

Evaluation and measurement should be considered together. Measurement answers the questions of how much, how many, and how often; it is concerned with quantities and qualities in evidence. Evaluation goes beyond the mechanics of testing and measuring by appraising the results of measurement in the light of aims and objectives. Thus, measurement and evaluation are related but not equivalent.

[1] Chapter 2 will relate the teacher's status to the aims of education and inquire into the true nature of educational progress.

In education, evaluation is a process of judging the effectiveness of educational experience. It answers the question of whether or not a particular experience has value. It is a continuous process. In its broadest sense it concerns the advancement of the total program and involves not only the data about the excellence of the school plant or pupils but also the way in which the school serves the community. Since evaluation is so necessary and important, all teachers need instruction to carry it out properly; but because no two teachers think and act exactly alike, evaluation practices vary. Certainly they are related to the teacher's own educational philosophy. Ross is of the opinion that evaluation requires the exercise of mature judgment.[2] Melchior holds that it is related to truth He stresses the fact that all teachers are truth seekers and says that to find the truth they must appraise their work carefully.[3] This is a noble idea that should challenge the average teacher and make him tremble a little when thinking of his responsibilities toward youth.

Historical Background in Measurement

A study of history indicates that, as man became more and more civilized, he became more exact in his measurement of human ability The body was appraised by the sculptor, soldier, and teacher alike They were interested in what the able and efficient body looked like The Athenians practiced general ability grouping rather than a rigid chronological age classification. In the days of Hippocrates (460 B C) and the stimulating days of the Romans, physical activity became almost a science, and there was much concern about the effect of diet and exercise on the total health of the body.

The Era of Scientific Measurement in Physical Education

The exact measurement and appraisal of man's physical structure and function began many years ago This was due chiefly to two influences. In the first place, the intellectual climate of the thirteenth century encouraged men like Roger Bacon to make exact observations of natural objects. Such observations eventually developed into the scientific method. The scientific method had an effect on teaching, causing it gradually to become more scientific. Experimentation with teaching methods and the content of the curriculum did much to show the need for careful measurement. This advanced the cause of evaluation, especially in education circles. The other influence on the early appraisal of human capacities was that physical and health educators were trained physicians. Anthropometric measures related to height, weight, body build, skeletal proportions, lung capacity, and

[2] Clay C. Ross, *Measurement in Today's Schools*, 3d ed, Englewood Cliffs, N.J., Prentice-Hall, Inc., 1954, p 488.

[3] William T. Melchoir, *Instructional Supervision*, Boston, D. C. Heath and Company, 1950, p. 23.

muscular strength were all measures of physical condition studied by early physical educators As far back as the middle of the nineteenth century, physicians vitally interested in youth were considering the relationship of these various anthropometric measurements to health status, motor skills, and physical performance in general Early physical standards were prepared in 1861 by Dr Edward Hitchcock, of Amherst College, and shortly thereafter by Dr Dudley Allen Sargent, of Harvard University.

President William A. Stearns, of Amherst College, greatly influenced American research and evaluation in physical education. He appointed Dr. Hitchcock to the chair of Hygiene and Physical Education, the first of its kind in America He thought that "human stature would increase as men practiced principles of good hygiene and exercise "[4] In order to become more familiar with anatomical measurements, Dr. Hitchcock corresponded with Galton, in London, and he began to gather data on his own freshman classes He checked "eight items of age, weight, height, chest girth, arm girth, forearm girth, lung capacity and pull, which were taken from every student from 1861 to about 1880 "[5]

On November 27, 1885, a group of American pioneers in physical education discussed the topic, "What is the best system of measurement " This was at the very first meeting of the American Association for the Advancement of Physical Education, held in Brooklyn, New York [6]

It was at this meeting that Dr Hitchcock recommended, on the basis of his twenty-five years of experience at Amherst College, that each student have a medical examination upon entrance to college Dr Sargent concurred in this recommendation and went on to say that "uniformity is needed to give value to statistics So as to measurement of muscle, of strength and stature. Capitalists and educators will respect aggregated influence more than the utterances of individuals and specialists All conditions of age, sex, mental and physical development must be studied "[7] Interest continued in the new professional association, and two years later, in 1887, Dr. Hitchcock asked the group assembled to take time in their teaching to make a careful study of how to find and point out remote and immediate weaknesses and tendencies to weakness in those who were "in our hands for study and advice "[8]

In the years that followed, Sargent practiced what he preached regarding the study of physical development. He measured muscular strength and endurance and related them to some forty-four anthropometric measurements

[4] Thomas LeDuc, *Piety and Intellect at Amherst College, 1865-1912,* New York, Columbia University Press, 1946, pp 129–130

[5] E. Hitchcock and H H Seelye, *An Anthropometric Manual,* Amherst, Mass., Press of Carpenter and Morehouse, 1893, p 5.

[6] *Proceedings of the Association for the Advancement of Physical Education,* Brooklyn, N.Y., Rome Brothers Steam Printers, 1885.

[7] *Ibid.,* p. 7.

[8] *Proceedings, Third Annual Meeting, American Association for the Advancement of Physical Education,* Brooklyn, N.Y., Nov. 25, 1887, p. 8.

in an attempt to understand and appraise his college men. In all his work he stressed the exact relationship existing between physical strength and the capacity for performance. He and his colleagues did much to advance the cause of scientific measurement and to spread the view that health is a positive characteristic. Much of the early research on motor ability and physical fitness carried on by men like Brace, Cozens, Rogers, McCloy, Cureton, Larson, and Clarke was influenced to a considerable degree by Sargent's basic findings.

Toward the end of the nineteenth century physiologists became more and more interested in physical condition and its bearing on physical education. The relationship between muscle functions and the condition of the circulatory mechanism was studied, and a number of cardiovascular tests appeared. It was not until about 1920, however, that a satisfactory cardiovascular test was employed to measure the response of the heart to physical exercise. This test, by E C Schneider, became famous during World War I when it was used to ascertain the physical fitness of young men for military duty [9]

The Era of Scientific Measurement in Health Education

While anthropologists, physicians, and physical educators were measuring the health status and physical performance of man, other groups of persons were concerned with the scientific aspects of health knowledge and health behavior. While Hitchcock was busy at Amherst, great discoveries were being made in Europe that pointed the way to public health The French bacteriologist, Pasteur, had made his discoveries, and Lister was busy carrying them out Men like Koch in Germany, with his study of organisms, and Metchnikoff in Russia, with his research on immunology, started people thinking about public health all over the world. Disraeli proclaimed public health the first duty of the statesman The way for school health education was prepared when eminent scientists such as Rudolph Virchow urged medical examinations of school children Other men, like William T. Sedgwick (1885 1921), brought about widespread acceptance of the importance of knowledge of sciences basic to environmental control of disease. By 1894, the Boston schools were ready to begin the first regular medical inspection programs in America. Then followed Chicago, New York, Philadelphia, and other cities From 1900 to 1915 a continued effort was made in the schools to appraise health status by giving medical examinations.

At first the teaching emphasis was on health knowledge as pure subject matter. It became apparent after awhile, however, that the educator must appraise the quantity and quality of the health information in terms of its effect on attitudes and habits of students. As far back as 1910 men like

[9] E. C. Schneider, "A Cardiovascular Rating as a Measure of Physical Fitness and Efficiency," *Journal of the American Medical Association,* 74:1506–1507, May 29, 1920.

Luther H. Gulick and R. Tait McKenzie stressed the need for appraisal in school health and hygiene.[10]

After World War I some headway was made in appraising school health programs. In 1922, Claire E Turner carried out his now famous Malden, Massachusetts, school health demonstration project. Over a two-year period he compared the children of three fourth, three fifth, and three sixth grades in two school buildings, with a control group of similar children in two other schools.[11] Health education was developed with the experimental group, and data were collected regarding growth, health status, and health habits. The demonstration was successful and did much to promote school health and instruction in the elementary grades.

Contemporary Interest in Evaluation

Since World War II, both professional educational conferences and citizens' committees have dealt more than ever before with the need for evaluation in the nation's schools Such questions as "Where are the schools headed?" and "What subject matter is important?" have been asked. Persons from all walks of life have become interested in such health and physical education topics as the physical fitness of youth, the value of intramural athletics, the relationship of exercise to morbidity and mortality from coronary heart disease, obesity as a disease of civilization, fluoridation of the drinking water, and recreation skills for leisure activity Professional and nonprofessional people alike point to the need for proper evaluation

Because President Dwight D. Eisenhower was disturbed over the results of a physical fitness survey, he called a special physical fitness conference [12] The purpose of the conference was to bring together a number of leading physical educators, physiologists, medical experts, and recreation leaders so that they might focus their attention on the kind of physical activity programs that build total fitness [13] The objective was *total* fitness. This conference was widely publicized and spurred a number of school personnel to test and measure their pupils for the first time In some places the fitness of children was considered on an individual basis for the first time.

More and more published reports emphasize the fact that there is a need for greater health attention, especially in rural areas and small towns where

[10] Luther H. Gulick, "Measurement as Applied to School Hygiene," *American Physical Education Review*, 16:239–241, April, 1911

[11] Claire E. Turner, C. Morley Sellery, and Sara Louise Smith, *School Health and Health Education*, St. Louis, The C V. Mosby Company, 1957, p 47.

[12] The fitness survey that prompted the President to call the conference was one conducted by Dr. Hans Kraus, of New York University, who compared European school children with American children along the Eastern seaboard and reported that European children were stronger and more physically fit.

[13] The American Association for Health, Physical Education, and Recreation, "The President's Conference on Fitness of American Youth," *Journal of Health, Physical Education and Recreation*, 27:8, September, 1956.

health services are deficient. The American Medical Association found that one of six small cities had no school nursing service; half these cities had no school physician. In many places a health examination is nothing more than a brief inspection by the classroom teacher Teachers of health and physical education in such situations are called on to screen pupils for referral to family physicians

At the Sixth National Conference on Physicians and Schools a number of recommendations were made relative to health and evaluation. These were made from the point of view of both the school physicians and the health education personnel One of the most enlightening points made was that the school physicians need help from the teachers in order properly to appraise a student's health status It was also pointed out that, when there is a high quality of school screening examinations, valid referrals result [14] This is a significant statement, for it suggests that if detection methods are inadequate, owing in part to uninterested or untrained teachers, students are not referred to the school health service at an early date It is the general opinion today that the success of a school health program is related to early detection and referral, plus a follow-up to secure essential medical care

Current philosophy and practice hold that physical education personnel have a definite responsibility for remedial or individual physical education; they must be able to appraise health status and work efficiently with the school and family physician Such efficiency requires continuous appraisal and subsequent modification of the professional curriculum to meet the students' needs In addition, schools are more and more concerned not only with defects and organic disturbances but also with mental and emotional conditions as shown in attitudes, personality traits, and habits

It is certainly true that since World War II a great amount of interest has been demonstrated in evaluation It is not uncommon for public school administrators to point to the need for measurement and evaluation as second only to the need for corrective physical education If the need is for more and/or improved corrective physical education, this, too, is significant. It is hardly possible to carry on a remedial program in physical education without determining individual status, adjusting the curriculum to the student, and noting improvement or lack of it This alone necessitates better understanding of measurement techniques and evaluation

One of the most important occurrences in recent years to have a direct bearing on evaluation in health, physical education, and recreation was the preparation of evaluation schedules for institutions that prepare teachers. The American Association for Health, Physical Education, and Recreation, cooperating with the American Association of Colleges for Teacher Education, prepared elaborate evaluation schedules for use by teacher-

[14] "Report of the Sixth National Conference on Physicians and Schools," *Journal of Health, Physical Education and Recreation*, 29:22–23, February, 1958.

education institutions offering a major program in health, physical education, and recreation.[15] Any college that prepares teachers can readily evaluate its health or physical education major program and identify strengths and weaknesses. This kind of appraisal leads to improved teacher-education programs. Specifically, it improves the content of measurement and evaluation courses.

Trends in Evaluation

Since 1930 the extensive use of measurement and standardized tests has improved education as a scientific process. The effect of validating the academic curriculum and methods through scientific inquiry has proved stimulating to medical and educational workers associated with health and physical education. As a result, much has been accomplished.

Some of the noteworthy trends in evaluation practices today, as compared with past practices in both health and physical education, are as follows:

A. Health Education Area

Past practices	*Present nationwide trends*
1. Physical examinations of school children were incomplete because of shortage of medical personnel and improper organization for appraisals	1. Physical examinations of school children are given in less haste and are more thorough, with better use made of the physician's time.
2. Screening examinations for hearing and vision losses were carried on haphazardly, with a good deal of subjective measurement.	2. Screening examinations for hearing and vision losses, together with weighing and measuring of pupils, are popular activities, carried on with objective measures.
3. Schools were short of medical personnel. Teachers, not prepared to appraise health status, did their best to detect abnormalities.	3. Teachers cooperate with school nurses, dental hygienists, and other medical personnel in an effort to determine the health status of the child. Many teachers have some training to improve their observation of health abnormalities
4. Schools did not know how to use physicians. They were concerned only with giving physical examination.	4. Schools have effective health service departments designed to evaluate the total school health effort.
5. Health committees or health councils were unheard of.	5. Health councils operate in a democratic way to determine health status and analyze health behavior
6. Defects and abnormalities that were revealed often were not followed up and corrected.	6. The objective is to evaluate the child as accurately as possible and then see that something is done about it.
7. Tests of health knowledge were designed to measure health information only.	7. Tests of health knowledge are improved, and there is increased use of tests designed to measure health attitudes and health habits.

[15] George F. Anderson, "Evaluating Teacher Education," *Journal of Health, Physical Education and Recreation*, 24:11–13, February, 1953. Copies of schedules are available from AAHPER; *Evaluation Schedules in Physical Education*, American Association for Health, Physical Education, and Recreation, 1201 16th St., N.W., Washington, 1952.

B Physical Education Area

Past practices	*Present nationwide trends*
1 Use of tests of physical fitness was limited to a few colleges	1 There is widespread use and an awakening interest in tests of physical fitness and measurement of the factors that constitute physical fitness
2 Motor skills were taught to increase neuromuscular coordination, but little attempt was made to measure this quality in physical education classes	2 There is a greater appraisal of motor skills, individually and in terms of general motor performance
3 Little effort was made to appraise and relate physical performance to total health, welfare, and scholarship	3 More and more efforts are made continually to relate physical performance and specific factors such as muscular strength and endurance to total health behavior, academic achievement, and personal happiness
4 Practically no attempt was made to measure objectively school progress in the development of specific sports skills	4 Appraisal of specific sports skills, in team, individual and recreational sports, is increasing
5 It was unheard of to take physical education class time to give written tests	5 The administration of short written tests pertaining to specific physical activities and physical education objectives is becoming more common

C Administration

1 Little effort was made to appraise the health and physical education program from an administrative viewpoint	1 Directors and supervisors of health and physical education attempt periodically to measure objectively the effectiveness of the program

Types of Measuring Devices

A variety of devices is available for health and physical education evaluation purposes Some are more useful than others

Observations. To the teacher who knows what to look for, observation can be a most fruitful evaluation technique If it is sometimes held in low repute, that is only because it tends to be subjective It need not be; observation can be quite objective, especially if the observer exercises self-discipline and is critical and precise Also, observations can be made more objective when guided by rating scales and check lists

Check Lists and Rating Scales. A simple list of things to look for when engaged in appraising pupils makes the difference between a thorough and objective evaluation and one of limited value

The check list or rating scale brings order to the observation It is usually put together after an analysis of pupil behavior in a given situation. It may be used, for example, to check the sanitary conditions of a building, to rate an athletic equipment supply room, to appraise individual social efficiency on the playground, or to measure the outcomes of a unit of work on a certain health topic (see Chapter 5). Some lists and scales are answered simply with a "yes" or "no."

Example

Is there a drinking fountain at each end of the building?
Yes________ No________

Other types are more carefully graded and are better for noting progress over a period of time

Example

Is there good sitting and standing posture?
Always Usually Seldom Never

Questionnaires. A series of specific questions can be listed for use in a particular health or physical education element A questionnaire is especially useful in eliciting activity, sleep, and food habits When used on an individual basis, especially in health education, they are helpful in eliciting the basis of personal health problems

Interviews. As an effective measuring device the interview falls somewhere between questionnaires and observations It can be quite effective in ascertaining how students feel about a program or how they themselves feel and why.

Records. In almost every school there are a number of vital health records available for teachers to look over Growth factors, height, weight, results of screening examinations, dental health status, and mental and emotional observations are often easy to find Cumulative and anecdotal records provide the teacher with usable accounts of pupil health behavior that may be related to health and physical education classwork and performance

Self-testing Activities. It is useful for pupils to appraise their own physical performance or health improvement What the pupil thinks of himself is fundamental It is an effective device for measuring attitudes Students rate themselves on such items as posture, sports skills, personal appearance, food habits, mental outlook, and attitude toward others on the playing fields (see Chapter 5)

Written and Oral Tests. Tests are the most common means of appraising the extent of pupil knowledge in use today Oral tests are excellent for determining what a pupil knows about a topic, but they are time-consuming Written tests, when properly constructed, can be used to appraise physical education skills, game knowledge and strategy, game rules, health understanding, health attitudes, and health habits.

Performance Tests. These are tests of individual skill, usually measures of motor ability and sports skills. There are also tests of physical capacity that involve a performance that is essentially unskilled, such as measures of strength, endurance, flexibility, balance, and reaction time.

Pupil Opinion Surveys. A survey is used by the teacher to ascertain the personal and public health interests of pupils. It is a means of identifying

pupil attitudes for the possible purpose of improving the instructional program; what students think and feel about physical education, health, or recreation can be surveyed and the information used in a number of ways.

The Case Study. The case study is a means of gathering specific and general information about a pupil for the purpose of health guidance or instruction. It is time-consuming, but in the hands of the health counselor it can be a most fruitful means of studying health behavior.

Measurement and Evaluation Applied

Interest in measurement and evaluation in health and physical education has grown in an orderly fashion Teachers and research workers have had ample opportunity to think about and explore the usefulness of specific tests and measures *The current emphasis is not too much on tests themselves as static things but on their application to problems in health and physical education.*

Today tests and measures are classified by their function This basis of classification is well over thirty-five years old.[16] Tests measuring aspects of posture are posture tests; those pertaining to health facts are called health knowledge tests Currently in the public schools and institutions of higher education properly classified tests are being employed more and more as devices to evaluate programs

Measurement, followed by careful appraisal of the findings, is generally considered as an administrative device to aid teachers in helping their pupils. The recorded results of measurement are sometimes stored away in school closets; these results, all too often, have not been applied They represent a certain amount of wasted effort, because unless the results are used no testing program is worth the time and energy it consumes To be purposeful, a measurement program in health and physical education must progressively eliminate chance and waste in health instruction and physical activity

Measurement devices and evaluation techniques are useful in numerous ways, so much so that the administration of the school is dependent on them. They are used in the following ways:

1. To appraise behavior, knowledge, attitudes, and appreciation
2. To classify pupils
3. To determine pupil, teacher, and program status
4. To measure program efficiency
5. To measure progress
6. To analyze students
7. To measure teacher efficiency
8. To measure pupil efficiency
9. To contribute to research

[16] David K. Brace, "The Classification of Tests of Physical Education," *American Physical Education Review*, 31:115–116, December, 1926.

They are especially useful for winning the support of the public and the administration for a new program, method, or facility. They are indispensable to the teaching staff in a school that has a well-defined purpose.

A Look to the Future

From all indications future promotion of learning will depend to a large extent on a continued scientific appraisal of school programs, together with corresponding modifications in curriculum content, methods of instruction, and pupil advancement.

Although contemporary interest in evaluation has been discussed and the chapters to follow show that many test materials and measurement devices are available today in health and physical education, there is still a noticeable lack of measurement and evaluation practices in many schools. Much of this is because a large number of health and physical education personnel are not sufficiently convinced of the need for proper evaluation, despite undergraduate and graduate college courses in the area. There remains, therefore, the challenge to make these college courses so meaningful that the graduate will be enthusiastic about measuring his progress toward educational goals. Also, in the years ahead better organized programs of in-service training will have a more beneficial effect on the professional activities of health teachers and physical education instructors.

The aim of this book is to help teachers realize that the more valid the measurement and the more judicious the evaluation of results, the sooner the goals of education may be reached.

SELECTED REFERENCES

American Association for Health, Physical Education, and Recreation *Research Methods Applied to Health, Physical Education and Recreation,* 2d ed., National Education Association, Washington, 1959, chap 1

Clarke, H. Harrison: *The Application of Measurement to Health and Physical Education,* Englewood Cliffs, N.J., Prentice-Hall, Inc., 1959, chap. 1.

Cowell, C. C: "Evaluation versus Measurement in Physical Education," *Journal of Health and Physical Education,* 12 499–501, November, 1941.

Haag, Jessie H.: *School Health Program,* New York, Henry Holt and Company, Inc., 1958, chap. 28.

Lamb, Arthur, Josephine L. Rathbone, and Peter V Karpovich. "Contributions of Physical Education to Measurement," *Journal of Health, Physical Education and Recreation,* 21:68–72, February, 1950

Ross, Clay C.: *Measurement in Today's Schools,* 3d ed., Englewood Cliffs, N.J., Prentice-Hall, Inc., 1954, chap. 1.

Willgoose, Carl E.: "Physical Fitness in Our Schools," *Education,* 76:508–509, April, 1956.

———: "Health Education and the Classroom Teacher," *Education,* 77:451–452, April, 1958.

CHAPTER 2

Aims and Objectives of Education

> The mariner, blown out of his course by adverse winds and sailing long under clouded heavens among dangerous reefs, seizes the first opportunity to get his bearings and charts his way by fixed marks of sky or land. Likewise, in the management of human affairs, although the analogy is not exact, it is often necessary for leaders of state, the professions, and callings, amid great disturbances, to take their reckonings—to recur to first principles.
>
> *Educational Policies Commission.*[1]

A sense of direction in education, as in any field of endeavor, depends on sound objectives. Failure to refer to these objectives from time to time causes teachers and administrators alike to flounder in the ever-broadening sphere of education. Unless considerable thought is first given to goals and purposes in education, measurement and evaluation mean very little.

More than one-fifth of all the people in America are enrolled in schools and colleges. This statement emphasizes the fact that education is big business. It is not conducted by chance by persons selected at random but by qualified groups of professional persons who know what they want and who have given primary consideration to the direction in which they are moving. Direction depends completely on basic concepts—the aims and objectives of the educators.

Educational Objectives

Maxwell Garnett points out that the aim of education cannot be defined until the aim of life has first been determined. The goal of human life, he says, is that human societies should aim at making progress and advancing steadily toward the fulfillment of some far-reaching purpose.[2] John Dewey would agree, but he would carry this view one step further.

> To profess to have an *aim* and then neglect the means of execution is self-delusion of the most dangerous sort. When we take ends without means, we degenerate into sentimentalism. In the ideal we fall back upon merely luck and chance and magic or exhortation and preaching.

[1] Educational Policies Commission, *Policies for Education in American Democracy*, National Education Association, Washington, 1946, p. 3.

[2] J. C. Maxwell Garnett, *Education and World Citizenship*, Cambridge, England, The University Press, 1921, p. 315.

Thus, aims suggest a "means of execution." Something has to happen. The word "education" bears this out, for it means "to lead forth or out." It refers to drawing out something latent or potential from a person. It suggests a change in some particular direction. Education, therefore, may be defined as a process of changing behavior toward certain preconceived goals.

Herbert Spencer offered a fivefold classification of aims that were essentially similar to modern goals.[3] For Spencer, education concerned (1) life and health, (2) earning a living, (3) family rearing, (4) citizenship, and (5) leisure. This list does not differ greatly from that of the *Cardinal Principles of Secondary Education.*[4] These principles are health, command of fundamental processes, worthy home membership, vocation, citizenship, worthy use of leisure, and ethical character. As far as health and physical education are concerned, both sets of purposes stress health as the primary aim and point to the need for education for leisure.

Probably the most far-reaching set of educational purposes is that put forth by the Educational Policies Commission of the National Education Association.[5] There are four categories of purpose: self-realization, human relationship, economic efficiency, and civil responsibility.

Health Education Objectives

There is agreement on the need for health education in the schools. Because health is so broad in definition, the ultimate scope of health education is great indeed [6] Health objectives, therefore, cut across several educational fields They are the responsibility of both the elementary school classroom teacher and the high school instructor They are basic to the activities of the school physician, nurse-teacher, dental hygienist, nutritionist, physical educator, and school psychologist.

Health education may be defined as "the sum of experiences which favorably influence practices, attitudes and knowledge relating to health."[7] In the American school system it is essentially a three-part process Each part has a specific program, which must be evaluated in terms of its own objectives. These parts of health education—health services, healthful school environment, and health instruction—were defined in 1934.[8]

[3] Herbert Spencer, *Education, Intellectual, Moral and Physical,* New York, Appleton-Century-Crofts, Inc, 1860

[4] Commission on the Reorganization of Secondary Education, *Cardinal Principles of Secondary Education,* Bureau of Education Bulletin 35, 1918

[5] Educational Policies Commission, *op cit*, p. 189

[6] The World Health Organization of the United Nations defines health as "a state of complete physical, mental and social well-being, and not merely the absence of disease and infirmity"

[7] This definition of health education is one agreed upon by the Joint Committee on Health Problems in Education of the National Education Association and the American Medical Association. Refer to *Health Education,* National Education Association, Washington, 1953.

[8] From the Committee Report, Health Education Section, American Physical Education Association, *Journal of Health and Physical Education,* December, 1934 p. 26.

Health services "comprise all those procedures designated to determine the health status of the child, to enlist his cooperation in health protection and maintenance, to inform parents of the defects that may be present, to prevent disease, and to correct remediable defects"

Healthful school living "designates the provision of a wholesome environment, the organization of a healthful school day, and the establishment of such pupil-teacher relationships as make a safe and sanitary school, favorable to the best development and living of pupils and teachers."

Health instruction "is that organization of learning experiences directed toward the development of favorable knowledges, attitudes and practices."

These definitions suggest some kind of continuous measurement and appraisal in order to be *certain* that the program is accomplishing its ends.

Physical Education Objectives

> Let us have men whose manhood is only the continuation of their boyhood, natural characters still, such are able for fertile and heroic action, and not that sad spectacle with which we are too often familiar, educated eyes in uneducated bodies.[8]

Many of the youth of the land lack the physical ability to participate in vigorous, growth-stimulating activities They are substrength in structure and in many cases deficient in basic skills, such as running, jumping, climbing, swimming, and throwing They simply have not been exposed to first-rate physical education activities It is almost an axiom that *the more complex civilization becomes, the less valid is instinctive behavior and therefore the greater is the need for education to preserve health*

When Emerson referred to "educated eyes in uneducated bodies," he was making a point in favor of fitness for living in the world of his day The civilization of today demands even greater attention to fitness. Exercise, nutrition, and rest build a quality of physical fitness that is a prerequisite for all other activities There is little opportunity for service—to be socially efficient or even to be charitable—unless one has the physical capacity to get started, keep going, and give of himself To think about great deeds but lack the physical capacity to perform them is indeed tragic It is physical capacity in addition to sound mental and emotional health that is the responsibility of the educator.

Physical education is education through physical means It is education primarily through large muscle activity, and it performs a unique function in the school program. It concerns total fitness: physical, mental and emotional, and social It involves fundamentally physical activity and physical skills, which not only add health and strength to the organism but provide it with the capacity to lead a rich and harmonious life in a close-knit society. It also concerns posture, general motor ability, and appropriate social behavior.

[8] R. W. Emerson, "Essay on Education," Boston, Houghton Mifflin Company, 1883.

Physical education challenges parents, classroom teachers, and specialists alike to teach and appraise the progress that boys and girls make in mastering knowledge and skills in physical activity. The physical activity program cannot be treated simply as free play or as a means to blow off steam. Research suggests that fitness in later life is, among other things, related to bona fide skills acquired during the early years.[10] The body remains functionally efficient for many years if there is regular exposure to physical activities, especially to those carried on during leisure moments in a pleasant atmosphere. Morbidity and mortality from coronary heart disease is twice as high among persons who do not lead a physically active life. Many diseases and defects associated with psychological stress increase wherever persons are not diverted from the fast pace of modern society.

There are at least four major objectives for physical education: physical fitness, social efficiency, culture, and recreational competency.

Physical Fitness. Physical fitness may be defined briefly as capacity for activity. It is a positive quality and is closely related to diet, exercise, rest, and emotional outlook. It is also a physical element of behavior that permits the person to be active. It can be referred to as organic vigor or vitality. It is demonstrated through physical performance. It appears in the person who radiates enthusiasm for living. Although related to health in general, it is more specific when carefully evaluated For example, several persons may be thoroughly checked by the school physician and found to be free from disease and defects, and yet they will vary in the degree to which they can perform physically Some will tire in walking a short distance Others may run the same distance without being winded The greater the physical fitness, the greater the physical endurance and precision of movement. The greater the physical fitness, the longer a person will be able to keep going; he will be able to perform more efficiently and at greater speeds and to recuperate faster from fatigue.

In performance the physically fit person has the following characteristics:

1. Strength enough to be ready for tasks encountered in everyday routine and in emergencies.
2. Stamina (endurance) to continue necessary tasks without undue fatigue, and energy enough to participate in recreational activities after a day's work.
3. Cardiorespiratory endurance for sustained effort in activities involving motion of the entire body.
4. Agility to be able to make a wide range of movement easily.
5. Speed to be able to move rapidly when personal safety demands it.
6. Control to coordinate body movements skillfully.

[10] Dr. Jay B. Nash has shown that more than 70 per cent of recreational activities of persons over the age of twenty-one were enjoyably indulged in before the age of twelve.

The objective of physical fitness is not the same as total fitness; but it is a major contributing element to total fitness and as such demands the full attention of educators—attention demonstrated by proper measurement and evaluation practices.

It may not be many years before a considerable body of evidence is available to support the premise that failure to develop physical fitness puts a limitation on genius *Mens sana in corpore sano* is not a new expression. The quality of the mind is made known through the competent actions of the body. Poor health and lowered physical capacity reduce one's *ability* to perform—as a genius or as a moron Terman recognized this years ago when he said that good things tend to go together. Man approaches his potential mental capacity only when he is capable of putting his thoughts into action.

Social Efficiency. Social efficiency involves human relationships It is the ability to get along with others and exhibit desirable standards of conduct. It is a kind of social well-being or social health School health education and physical education contribute measurably to this objective by providing learning situations in which pupils can express themselves through self-restraint, initiative, perseverance, courage, cooperation, leadership, followership, and loyalty to the group Good health by itself supports social efficiency, for without the physical capacity to pursue social goals, a person is handicapped and less able to do for himself or others

Physical education, perhaps more than any other special subject in the curriculum, is purposely organized to deal specifically with the elements of sound social behavior Through games and dances pupils are exposed and observed as they assume their responsibilities when working with others. Cooperation is born out of a feeling for others, or group consciousness. From awareness of others develops a degree of participation, which is demonstrated through followership and leadership In a play situation the essential characteristics of social efficiency are practiced through sportsmanship. In fact, sports-minded people have demonstrated the meaning of this word so well that it is used continually outside the world of sports. Over the centuries sports have left their mark of value Justice and honesty, terms that the philosopher may analyze at length, are readily described by physical education personnel as the essential elements of fair play. If it is a fair situation, it is a sporting one, for the very essence of sportsmanship is fair play. When activity groups and teams are purposely set up so that none has an unfair advantage and when equality between players is the rule, then there is opportunity for sportsmanship in play, with the corresponding development of personal adjustment and optimum group behavior.

Physical education has much to offer in the realm of social behavior, but providing an opportunity for the development of social efficiency is not enough. Measurement techniques need to be applied in order that teams can be equalized more scientifically, that the degree of such desirable ele-

ments as cooperation and leadership can be evaluated in each child, and that physical education will be seen clearly as a moral force in present-day society. There is little doubt that the influence of a sound physical education program is great indeed. Western civilization appears to some people to be gradually slipping into moral chaos. Certainly the child's life in the gymnasium, on the playground, and in the pool is rich with opportunities for carefully cultivating moral values.[11] If the teacher of physical education is unaware of the potential influence of the program and fails to teach and appraise moral values as strenuously as he teaches skill, then physical education will miss its greatest chance to serve humanity.

Culture. Perhaps the least considered objective of physical education is the development of the cultured person Culture is not something ethereal or celestial that cannot be grasped It involves a deep appreciation for life's activities. It is associated with the rich and full life.

Teachers of physical education who develop appreciation for rhythm and music through specific sports skills and dances and who employ form and color in creative activities are helping to develop the cultured person But there is more to it than this Every teacher of physical education who successfully teaches a pupil a *useful motor skill* is contributing to the cultural objective.

This is a much deeper expression of culture—an inseparable part of the physically educated person. The very meaning of culture substantiates this.[12] In general, the mark of a cultured person is refinement in mental and moral powers as a result of a particular training and enlightenment. The cultured person feels a concern for all things Careful cultivation, the product of learning activity that enriches and fills him with appreciation, and his attitude toward all the arts show a person to be cultured The part physical educators play in this process frequently goes unnoticed.

The physically educated person possesses, in addition to the qualities mentioned above, a kinesthetic sense or consciousness within the body structure for specific skilled movements This artistic experience manifested through kinesthetic sense appreciation is comparable to the appreciation shown toward the fine arts, literature, and music.

The accomplishment of difficult skills, and the gradual learning of new skills, involve the same nervous pathways and end in a neural pattern no less important to artistic appreciation than stimuli received through the smaller muscle groups. A man listening to a symphony, upon hearing the music, may unconsciously close his eyes and relax his body in the fullest appreciation. A man who has completed the exercise of skills involved in

[11] See Delbert Oberteuffer, "Strengthening Moral and Spiritual Values," in *Children in Focus*, Yearbook of American Association for Health, Physical Education, and Recreation, Washington, 1954, chap. 13.

[12] For a fascinating discussion along these lines, see the philosophical approach presented by Eleanor Metheny and Lois Ellfeldt, "Movement and Meaning: Development and Theory," *The Research Quarterly*, 29:274–276, October, 1958.

the daily workout, handball game, tennis tournament, or golf match and who has taken his shower and is ready to leave the place of activity will often unconsciously throw out his chest and stand a little taller, realizing a great feeling of satisfaction. The reactions of both the music lover and the sportsman are the outward manifestations of the cultured person. Both have a fine artistic sense; the former through small muscle sensorimotor response and the latter predominantly through large muscle sensorimotor response.

Recreational Competency. Ancient philosophers have stressed the attainment of a goal for which present-day educators strive: the worthy use of leisure time. The competent person is one who engages in recreational activities as a form of relaxation and diversion for the purposes of personal enrichment.

Man has come to realize that his survival is closely linked to a fine relationship between work and play. Civilization has its roots in hard work, but it flowers in leisure. During leisure, man becomes rejuvenated and revitalized; he discovers, creates, enjoys, and lives fully. Recreation is both physical and mental. It is rewarding because it relieves the frustrations, problems, and tensions of society as well as the numerous diseases of physical and psychological stress.

Recreational skills, many of which are physical, need to be mastered early so that a person can participate in various activities during a lifetime. It is as necessary to get away from work routine as it is to secure proper food and shelter. Recreational skills, therefore, need to be evaluated. It is important also that recreational skills be used with the right attitude. People are hardly educated for recreational activity if they work as hard at their play as they work at a given occupation. There is little true relief for the higher nerve centers when the same attitude prevails toward work and leisure. Pent-up emotions need to be let out, and glands of internal secretion need to rest from constant stimulation.[18] It simply is not enough to divert attention from work to golf; attitudes must also be diverted.

It is both interesting and challenging to note that history shows that a civilization declines when a great amount of leisure is available. Moreover, it is not uncommon for present-day philosophers to suggest that if man does not destroy himself in the space age, he will destroy himself through leisure. Such expressions of doom can only cause health and physical education personnel to evaluate more carefully what is taught in the name of recreation. There is more to recreational competency than meets the eye.

The Essence of Progress

Seneca once said that the greatest part of progress is the *desire* to progress. Desire is the key word. One's heart must be in the right place. When this is true, objectives mean something and there is a personal striving to make

[18] See Hans Selye, *The Stress of Life*, New York, McGraw-Hill Book Company, Inc., 1956; also Jay B. Nash, *The Philosophy of Recreation and Leisure*, St. Louis, The C. V. Mosby Company, 1953.

progress in a definite direction. When the desire is strong, the earnest teacher becomes frustrated and questions his own action. It is through measurement that he finds out just where he is. He determines whether he is running in circles or really progressing toward his goals.

The objectives of health education and physical education embrace so much that is closely related to the over-all purpose of education that teachers and supervisors in this special area might well be challenged at the opportunities available to modify favorably the behavior of their charges. All conscientious teachers must search periodically for answers to these questions: "Am I really changing behavior according to plan?" "How do I know?" "Can I trust my appraisal and evaluation techniques?" "How shall I improve my curriculum, methods and materials?" "Am I making progress?"

SELECTED REFERENCES

Bair, Don E.: "Identification of Some Philosophical Beliefs of Influential Leaders in American Physical Education," *The Research Quarterly*, 28 315–317, December, 1957.

Bucher, Charles: *Foundations of Physical Education*, St. Louis, The C. V. Mosby Company, 1952, chap. 1.

Jack, Harold K.: "Interpreting Health and Physical Education," *Journal of Health, Physical Education and Recreation*, 29.33–34, November, 1958.

McCloy, Charles H.: *Philosophical Basis for Physical Education*, New York, Appleton-Century-Crofts, Inc, 1940, p 30.

Morris, Van Cleve: "Physical Education and the Philosophy of Education," *Journal of Health, Physical Education and Recreation*, 27:21-23, March, 1956.

Nash, Jay B.: *The Philosophy of Recreation and Leisure*, St. Louis, The C. V. Mosby Company, 1953.

Nixon, Eugene W., and Frederick W. Cozens: *An Introduction to Physical Education*, Philadelphia, W. B. Saunders Company, 1959.

Oberteuffer, Delbert: "A Decalogue of Principles," *Journal of Health and Physical Education*, 18:12–14, January, 1947.

Price, Hartly D.: "The Establishment of Principles Which Are Essential for the Realization of the Objectives of Physical Education," *The Research Quarterly*, 19:92 93, May, 1948.

Rogers, Frederick R.: *Educational Objectives of Physical Activity*, New York, A. S. Barnes and Company, 1931, p. 26.

Steinhaus, Arthur: "Contributions of Physical Education," *The Research Quar-* [illegible]

Willgoose, Carl E.: "Culture and the Kinesthetic Sense," *Journal of Health and Physical Education*, 18:149-150, March, 1947.

———: "Health, the Fundamental Objective," *Education*, 68·1-3, April, 1948

———: "Physical Fitness: Our Primary Objective," *Journal of Health, Physical Education and Recreation*, 30:32-33, November, 1959.

Williams, Jesse F., Clifford L. Brownell, and Elmon L. Vernier: *The Principles of Physical Education*, 8th ed., Philadelphia, W. B. Saunders Company, 1958.

CHAPTER 3

The Criteria for Proper Measures

> Grant me the strength, time and opportunity always to correct what I have acquired, always extend its domain, for knowledge is immense and the spirit of man can extend indefinitely to enrich itself daily with new requirements. Today he can discover his errors of yesterday and tomorrow he may obtain light on what he thinks himself sure today.
>
> *Maimonides*

It has been said that orderly progress toward goals in education depends on measurement, because measurement and purposeful activity are inseparable. Generally, the more valid the measurement and the more careful the evaluation of results the sooner the goal will be attained. This becomes especially true if proper criteria have been established for the instrument of measurement.

Standards of Judgment

When the scientist becomes critical he expresses a reasoned opinion, involving a judgment of value or truth or an appreciation of beauty or technique. Such a critical observation is based on some standard of judgment. A rule or test or measure is usually referred to, and this may be called a *criterion*. A criterion is a standard of judging; without a criterion any judgment is little more than a guess.

In order for the student of measurement in health and physical education to judge whether some measures are better than others, standards of judgment are needed. What are some of the commonly accepted criteria of good measurement? What are some of the important characteristics of worthwhile tests? Can a test be used in certain circumstances? How much weight should be given to the results?

Criteria that are applied to measures in health and physical education do not differ from those applied in other areas of education. These include validity, reliability, objectivity, economy, norms, and interest.

Validity

A test that agrees with truth is valid. It measures what it is supposed to measure. If a test is supposed to measure health status, test results and

health status correlate. The thermometer measures temperature; the barometer, atmospheric pressure. The thermometer is not a valid measure of atmospheric pressure, but it is valid as a measure of temperature. It is valid as a health measure because readings of 101, 102, or 103°F almost always accompany highly disturbed physiological conditions. The two conditions correlate. There is a high degree of agreement between temperature changes and health status.

A test is valid for clerical aptitude if scores correlate well with later clerical proficiency. Likewise, a test is valid for swimming aptitude if the results match later swimming performance in a race. School marks, for example, are useful to note pupil progress, but they are only slightly valid indicators of achievement. Why is this true? Unfortunately, this is true because a number of pupils with low grades succeed very well after completing their formal education. This does not mean that grades should be discontinued. Rather, tests should be modified so that grades represent more truly the pupil's total behavior and potential, thereby becoming more meaningful.

Validity is a useless term until it is attached to an aim. "Valid for what?" is the question to be asked. Validity is quite specific. A written test or mechanical instrument is generally valid for one measurement. Very often when it attempts to embrace several items it loses some of its value.

In order to check its validity and to establish the *criterion of validity* itself, a test is frequently compared with another test of the same factor in which results are already known. This is sometimes difficult to do, especially if the test in question has been prepared to measure something that has never been accurately measured before. The foot ruler can be used to validate the yardstick, and the yardstick, in turn, can be used to validate the steel tape. Therefore, the validity of the steel tape can be checked by using the well-known yardstick as the *criterion*. In measuring a long room the steel tape is a more valid measuring instrument than the other two, primarily because it is more accurate. Extending the steel tape to its full length permits less error than moving the yardstick several times.

Dyer validated her test of general tennis ability against round-robin play and found a high correlation between the two.[1] Larson validated his indoor and outdoor tests of motor ability against some twenty-five separate motor ability items.[2] The test proved valid because persons who scored well on individual motor skills also scored well on Larson's test.

Sometimes the validity criterion is established by checking different groups. For example, the upper quarter of the group tested might be compared with the lower quarter. When McCurdy and Larson constructed their

[1] Joanna T. Dyer, "The Backboard Test of Tennis Ability," *The Research Quarterly* (Supplement), 6:63–65, March, 1935.

[2] Leonard A. Larson, "A Factor Analysis of Motor Ability Variables and Tests for College Men," *The Research Quarterly*, 12:98–100, October, 1941.

organic efficiency test, they found that the test discriminated satisfactorily between (1) a well-conditioned group of varsity 440-yard swimmers, (2) a poorly conditioned group of college infirmary patients, and (3) an average group of college freshmen.[3] The subjects used were the following:

1. Sixty college varsity swimmers and forty American Olympic swimmers in the peak of condition. This was the well-conditioned physiological group.
2. One hundred thirty-eight college infirmary patients examined soon after confinement in the infirmary for two or more days with respiratory disturbances. This was the poorly conditioned group
3. Two hundred eighty-six college freshmen This was the average group.

The biserial correlation between the test scores and the two physiological groups was .70 The correlation between the test and the 440-yard swimmers was .68.[4]

The combined judgment of experts is sometimes used to validate a test. This is generally done when it is difficult to find a similar test to use for comparison For example, a health knowledge test might appear to be most effective to the author of the test To validate this particular measure a group of competent judges might rate the test questions as satisfactory or otherwise, according to their respective judgment Highly qualified judges would be most helpful in this instance The procedure is time-consuming, but it often proves more beneficial than comparing a new test with several older ones of questionable value

Quite often tests constructed at the local level are commendable and help the teacher determine progress in many ways Frequently, however, their validity is never properly checked, and they are misleading Tests that mislead are far worse than no tests at all Homemade measures of health attitudes and sports skills sometimes fall in this category Suppose that a test called a skill test in soccer is given to 200 ninth-grade pupils In examining the test scores it is found that pupils who are poor soccer team players in many cases have test scores as high as those of good soccer team players. If it is further discovered that those who are just average team players receive high, medium, and low test scores in almost equal numbers, and that about one-third of the really good players receive good scores and

[3] J H McCurdy and L A Larson, "Measurement of Organic Efficiency for the Prediction of Physical Condition, *The Research Quarterly* (Supplement), 6:11–14, May, 1935.

[4] Generally, a coefficient of correlation over .80 for validity is a fairly good indication that the test measures what it is supposed to measure When the coefficient is down around .60, it may be considered only about 20 per cent better than chance. To convert a coefficient of correlation to a per cent better than a chance figure, use the predictive index formula:

$$PI = 100(1 - \sqrt{1 - r^2})$$

For an explanation of coefficient of correlation, see section in Chap 18. See Glossary for explanation of biserial correlation.

one-third receive low scores, one would be forced to conclude that this soccer skill test does not measure ability to play soccer. In this situation, however, the test might measure something else closely allied to success in soccer, such as running ability.

Validity is only as good as the other criteria of measurement. It rises and falls with such items as the reliability of the test, the interest of the measured and measurer alike, the quality of the instruments, and the length of the test. One should not put blind faith in tests of high validity, even when the coefficient of correlation is .90 or better, for a number of highly valid tests have proved inaccurate when judged by other criteria. A test, for example, that takes a long time to administer can be worthless if the testee and tester are disturbed by its length and rush through the items in an erratic or sloppy manner.

Reliability

Simply stated, reliability is trustworthiness. Will the test repeat itself? Is it consistent? Can the measurer, using the same process and measuring device, obtain the same score twice, three times, etc.?

Reliability is a relative thing; no two measurements of anything can ever be the same. There probably is no such thing as an absolutely exact duplication of anything. The passing of time inevitably results in changes. Conditions of the moment are never the same as they were the moment before. It is overly optimistic and impractical, therefore, to expect a test to be completely free from error. However, it is possible to obtain reliability in testing or carrying on research with a very small percentage of error. This can be done by selecting accurate instruments for the particular measurement. The foot ruler is valid for measuring distance, but triangulation by surveying instruments may be more valid because it is more precise. The stop watch measures time intervals more precisely than a dollar watch. Both are valid time recorders, but one is more reliable for a specific purpose.

Reliability is expressed in terms of error. An instrument is considered to have a probable error of a certain percentage. The problem is to determine how much error can be tolerated and yet maintain a satisfactory level of reliability. An error of 5 per cent in a manuometer measuring grip strength would not be excessive, but the same percentage of error in a watch would render it almost worthless.

The coefficient of reliability of a test or measure indicates how well it will repeat itself and whether it is accurate enough to be used. When the test reliability coefficient is less than .90, the test is too inaccurate for research purposes. Thus reliability is closely related to validity.

In addition to improving test reliability by choosing a more accurate evaluation instrument, one should consider the nature of the subject being measured. If it is a school child, care should be taken to see that he is feeling well. Often a pupil does not respond to a test of unquestionable validity because a physical or emotional disturbance has temporarily upset

him. Also, one should see that he is properly motivated and that he is generally prepared to be tested. Test administration, therefore, involves the proper orientation of the testee before the test is given. Uninterested and uncooperative persons influence the reliability of any measure.

Closely related to this is another administrative item—the measurement process itself. Is it complicated? Is it apt to be confusing? Is it lengthy to the point of being burdensome? Simplicity is a highly desirable factor in any test. There is always a danger, of course, of making a measure too simple, with the result that the ensuing test scores have little real value.

Finally, the nature of the evaluator is a factor in test reliability. He must believe in the worth of what he is doing, be skilled in the techniques, and register at least a mild enthusiasm for the whole process. In fact, it is this enthusiasm to find out that gives stature to the scientist, the competent teacher, and anyone else determined to progress toward some personal goal. The *attitude* one has toward evaluation in health education or physical education may have more to do with reliability and objectivity of a measure than almost any other single factor. It isn't *what* you do; it's the *way* you do it that counts. Just as the teacher must have faith in the worth of teaching, he must also believe in the processes of evaluation. In the long run his natural curiosity, coupled with a dedication to effective teaching, will result in improved appraisal technique and a corresponding improvement in child behavior.

Objectivity

Tests must be so constructed that equally competent persons can use them and get the same results. Objectivity can be described as the degree of uniformity between persons scoring the same test. It refers to the measuring processes and instruments used that are administered by different testers. Testers A and B agree with each other: the measure has objectivity. Objectivity is just as relative as reliability, for one would hardly expect several testers to arrive at exactly the same results.

Objectivity of a measure may be improved by making the test easier to administer. Sometimes this can be accomplished by rearranging a few test items or making the testing atmosphere a little more pleasant for those giving the tests. One should probably compute the objectivity of test scores given by partially trained people.[5] Helping these people improve their testing technique may be all that is needed to raise test objectivity. Standards for such coefficients are usually as high or higher than those standards expected on reliability.

Norms and the Concept of Normal

There are certain fundamental likenesses that all people have in common. People are never so different that they are not recognizable as human beings,

[5] Objectivity coefficients may be found by correlating the scores of the various judges. See Chap. 18 for information on how this is done.

differing only within the limits that human ranges allow. Yet every person differs enough that no two are alike.

What is normal? This question is often disturbing, for in every direction there abounds a multitude of things that are not as they are supposed to be. A man shuffling along with size 13 shoes catches our attention not when he towers well over 6 feet in height but when he is only 5 feet tall. His health may be perfectly normal, but the size of his feet is not. His feet are too long for a man of his limited height. An appropriate norm table would show this, for a norm is a kind of average of all extremes of existing characteristics.

More specifically, norms are standard points of reference that provide a basis for judgment. They are used continually by the conscientious educator and research worker because the processes of measuring involve *comparisons*. When a boy is taller than the average of the class, one immediately inquires into the average height of the other pupils. The average is standard for the class; thus, the boy in question is above the class norm. In the same way, if the average boy of a certain age, height, and weight executes a standing broad jump of 7 feet 2 inches, then 7 feet 2 inches is a norm against which to compare other boys of the same age, height, and weight.

Norms are temporary standards subject to revision. They depend on standard and average deviations of the test scores. They provide considerable value to all measures employed in education. Without them, the results of the most valid measures cannot be properly weighed; only simple comparisons can be made.

The norm is in no way an ideal. It is concerned with characteristics that happen to be, whether good or bad. It must be admitted that it is almost normal to see people with defective teeth, digestive upsets, nearsightedness, baldness, aching feet, advanced obesity, mental instability, etc. It even seems normal for close to a million people to die of heart disease each year. Yet these items can hardly be accepted as desirable even if they are within the normal range of what might be found in a cross section of the population.

Several persons may be alike (almost alike) in a certain characteristic yet vary immensely in another. In this instance it is difficult to say who is normal. One can only say that a person is normal with respect to a definite characteristic. For this reason most norm tables do not embrace several characteristics. Roger Williams has said that "if 95 per cent of people are normal with respect to one comparable item, only 90 per cent would be normal with respect to two items, 60 per cent would be normal with respect to ten items, and only a little better than half of 1 per cent when we include 100 items."[6]

The *use of norms* is sometimes misunderstood. The pseudo-scientist may be so wrapped up in scoring a student against his own performance that he overlooks the opportunity to compare him carefully with others. It is only

[6] Roger J. Williams, "Chemistry Makes the Man," *The Saturday Review*, Apr. 6, 1957, p. 42.

by comparing him with others that status can be accurately determined. Social status, physical fitness status, and program status are all more precisely explained with the establishment of norms. The more representative the norms, the more useful they are in ascertaining status.

In health and physical education, norms are often based on various combinations of age, height, and weight. With such norms test scores are adequately interpreted. In the school nurse's office one asks if a weight of 120 pounds is satisfactory. In the gymnasium the instructor wonders about the strength of the upper extremity when a boy performs 15 pull-ups. Is a strength index of 1500 good? These questions can only be answered when additional information is available. A weight of 120 pounds is satisfactory for certain persons; for some it indicates a serious underweight condition; for others it may relate to chronic fatigue; and for still others it may mean obesity. With norm charts based upon sex, age, height, and general body build, or other factors, a satisfactory answer can readily be reached. The same reasoning applies to the strength index. A strength index of 1500 may indicate adequate strength, great strength, or general weakness, depending upon the sex, age, body type, and weight of the subject. The 15 pull-ups take on greater meaning when it is discovered that this performance has a T-scale score of 50, representing the average performance of U.S. Navy personnel.

The whole argument for norm tables is based on individual differences. Present-day educational philosophies generally agree that these differences should be understood and identified. Curricula and methods, therefore, are appropriately adapted to individual differences. It is these differences that are recognized when norms are employed. Consider the following figures, which were obtained by J. Stuart Wickens[7]:

Test	Mean	Standard deviation	Low-high scores	Range
Skeletal age, months	161	12.5	122–184	62
Body weight, kg	48	10.3	34–70	36
Standing height, cm	158	7.7	142–174	30
Chest girth, cm	79	7.1	70–97	27
Lung capacity, cu cm	3,116	570.0	2,098–4,162	2,064
Leg lift, kg	396	81.3	192–576	384
Push-ups, parallel bars	5	3.4	1–16	15
Strength index	1,425	277.9	944–2,242	1,298
Standing broad jump, cm	169	17.4	127–216	89

These selected data indicate the great individual differences among boys of the same chronological age (thirteen years). Realistically these figures are

[7] From a doctoral dissertation by J. Stuart Wickens, University of Oregon, involving an analysis of maturity, structure, muscular strength, and motor ability of boys nine through fifteen years of age in the Medford, Ore., public schools.

norms; at least they are the data from which formal scales are constructed. Generally, norms contain many gradations of the range of scores (note norm tables in Appendix A). Frequently the mean represents the center value on the scale, and the gradations are in standard deviations from the mean. Certainly there is little doubt of the value of such information as this in evaluating the status of an individual child or an educational program.

Economy

Time is the essence of economy, and expense is close behind. In education, as in almost every other field of endeavor, time and money strongly influence the use of certain measurements.

An educational test may be quite desirable when considered in the light of the essential criteria already discussed, but the use to which the results can be put may not justify the effort and money involved in it. It may take a long time to administer, or the training period for testers may be a long one. Also, a number of diagnostic measures require considerable time to score accurately. Certainly the teacher of health or physical education must decide at an early date how much of the total class time will be used for instruction and how much for measurement and evaluation. Arriving at a proper balance involves the criterion of economy.

Several tests in physical education have an excellent coefficient of validity. The Illinois Motor Fitness Test is one. In this test, however, the utility factor is greatly reduced because one hour is needed to test ten students even when experienced testers are employed.

Interest

When pupil capacities are measured, will power and interest are involved. Some pupils have a natural desire to test their knowledge, skill, or fitness. Very often interest is developed psychologically by a program of motivation. In educational measurement and evaluation extrinsic rewards and competition are not used to get pupils to perform their best. The safest motivation of interest stems from an understanding of why the tests are being given. The appropriate explanation, therefore, boosts the pupils' interest so they voluntarily give their best performance. This makes the criterion of interest closely allied to reliability.

Interest in testing is further maintained when pupils know how the test results will be used. In hundreds of schools in this country youngsters in physical education classes know that fitness scores are being used specifically to improve their posture, raise their strength, and help them live a better life at the moment. They take their testing seriously. The same is true in a class where the results of a health knowledge test are discussed and used as a part of a subsequent lesson. *Good teaching generally combines a nice balance of appraisal activities with instruction. The two go hand in hand; they are not separate entities. Both are continuous processes.*

Test Selection and Evaluation

Evaluation goes beyond the mechanics of testing and measuring by appraising the measurements in the light of preconceived aims and objectives. Here is where some programs fall short. There are schools that do a good job of testing and measuring only to file the results away where they cannot be appraised and used. There are schools, on the other hand, that make a definite attempt to evaluate programs but with very little objective measurement. It is obvious that weaknesses exist in both the examples. What is needed in a school is a scientifically conceived program that measures objectively all possible factors and results in a strong follow-up plan.

In health and physical education there has been much discussion of the respective values of the various tests. The ability to select proper tests and measures appears to be one of the earmarks of the competent health teacher or physical education instructor. Selecting the best test, however, is not easy. Usually one must ascertain whether or not the test in question really meets the criteria of sound measurement. For this, the test must be examined carefully and tried out.

Tests are used only for specific purposes Whether or not a test is a valid and useful one for the purpose intended often depends on the professional opinion of research and evaluation experts in health and physical education. These experts have studied the test, used it, and compared it with others of the same nature Even in a book like this it would be impractical to cover the dozens of existing tests in the special areas. This text is primarily concerned with measures that (1) have the support of professional workers in a particular area, (2) meet, in general, all the criteria of proper measurement, and (3) have been tried in experimental application and verified by repeated application.

For some factors it is difficult to suggest a highly valid measurement instrument or technique; evaluation of these matters is limited accordingly. Better tests are needed for many items; yet it is quite possible to carry on a program of continuous appraisal with present tests and measurement devices and ascertain whether advancement is being made.

Numerous tests appearing in the literature have been prepared without proper attention to validation, have obvious limitations, and are of questionable utility. In an effort to evaluate a large number of the tests in health and physical education, the National Research Council of the American Association for Health, Physical Education, and Recreation published a report in 1950 and 1960 entitled *Measurement and Evaluation Materials in Health, Physical Education and Recreation.*[8] Information on criteria was gathered from original sources and subsequent research. The only tests

[8] *Measurement and Evaluation Materials in Health, Physical Education and Recreation,* National Research Council, American Association for Health, Physical Education, and Recreation, Washington, 1960.

included in the report were those that had been tried and verified. The degree to which specific tests were limited was indicated by professional workers familiar with the subject matter of the particular test. Also, numerical coefficients of validity were listed for each test so that the reader could understand the test value and act accordingly.

In selecting a test for a given situation, one must determine at the outset exactly *how the test results will be used.* Will the results be used to predict specific behavior? Will they be used to predict behavior in a general way? Is the test to be selected chiefly as a screening device? The audiometer, for example, is a very practical device for screening a group of school children for hearing loss; for this purpose it is a fine instrument But in the hands of the nurse or teacher it is not intended to take the place of an individual examination by a hearing specialist In this case the test selected has enough validity for *one purpose,* and its limitations are acknowledged There are many such useful tests. The Harvard Step Test, which is concerned with cardiorespiratory efficiency, varies in validity from 20 to 50, probably because the initial pulse rates are sometimes unreliable Yet, as a screening device, this test has value and has been used for many years A number of tests of health knowledge, understanding, attitudes, and sports skills have low validity, but they have utility because they give hints to student status or behavior They are not used as a basis for predicting future outcomes They simply give some meaning to observations and are sometimes improved as time goes on

Occasionally, a test maker will claim that his test measures students at both junior and senior high school levels, but on close examination it is discovered that the test was validated on the performance of senior high school students only. This does not mean the test cannot be used for both groups; it does mean, however, that when the junior high school pupils are measured, only a small amount of weight should be placed on the results until the behavior of a fair sample of junior high school pupils has been compared with that of the older group

Testing and evaluating is not easy work Continual scrutiny of all available tests is necessary in order to do the best possible work in the school system Some of the limitations of tests presently in use in health and physical education are as follows:

A screening device only.
Careful training of examiners necessary
No evidence of validity on elementary school level.
Time-consuming.
Norms based on limited sample.
Data for college men only.
Useful for boys only.
Norms not available.

Reliability information incomplete.
Requires expensive photographic equipment.
Limited to beginning classes.
Test constructed by the empirical method.
Considerable equipment needed.

Standardized Tests

The term *standardized tests* has been used for measuring instruments that have reached widely varying degrees of refinement through scientific procedures. There are a number of such tests of health knowledge, physical fitness, general motor ability, and sports skills. Most physical education and health teachers will be interested to evaluate the extent of standardization that has been done on the tests they use. The tests they construct for their own purposes will usually not reach an advanced stage of refinement until after several years of use. Few of them will become in a true sense thoroughly standardized. With improvement every attempt possible should be made to meet the criteria of a proper test: validity, reliability, objectivity, norms, economy, and interest.

It has been pointed out by Bean that the truly standardized test requires that four essential conditions be met.[9] The first condition is that test instructions must be clear and include specific details for keeping test conditions uniform. A test is weak when too much is left to the examiner's judgment. The second condition essential to standardization is uniformity of scoring. The third condition is adequate norms. This is worth referring to again for it seems often to be a subject of controversy. Certainly it is not always easy to prove that a representative sample has been secured. But makers of standardized tests attempt to provide norm tables even if they are employed only for simple comparisons. The fourth condition is that there must be necessary supporting research to establish such criteria as validity, reliability, and the level of test difficulty.

The teacher or administrator has a right to expect that accurate information will be furnished him by the author or publisher of any test. A number of valuable sources of evidence about health, physical education, and related tests are found in the current copy of *The Mental Measurement Yearbook* in new editions at frequent intervals.[10]

Item Analysis

Up to now we have been concerned with the *external* relations of a test; that is, how the results agreed with an independent criterion. Had we pursued it further in this chapter, it would have involved statistics and

[9] Kenneth L. Bean, *Construction of Educational and Personnel Tests*, New York, McGraw-Hill Book Company, Inc., 1953, p. 174.

[10] Oscar K. Buros (Ed.), *The Fourth Mental Measurement Yearbook*, Highland Park, N.J., Gryphon Press, 1953.

might have been called a form of statistical analysis. At the moment the purpose is to examine the *internal* properties of the test results, such as level of difficulty and discriminating power, both of which involve reliability.[11]

In all tests, whether physical performance tests or written tests, it is the hope of the instructor that they will be of the appropriate level of difficulty for the group. Thus, a test, to be valid, must discriminate accurately between good and poor students. For example, fewer persons in the upper third of a class should fail items that are failed by the lower third of the class if the questions are discriminating ones. In addition to this, attention should be given to the degree of difficulty of the question. If a golf skill item or health test item is so difficult that all pupils miss it, or so easy that all pupils pass it, there is no evidence of discrimination. Thus, after a written test is given, several questions may have to be revised. New skills may need to be worked out for a performance test.

The *difficulty rating* is used primarily with written tests. The rating of a question is arrived at by dividing the number of correct responses by the number of pupils answering the question multipled by 100. The lower the score the more difficult the question.

Example

In a baseball knowledge test 65 out of 110 boys answered question 3 correctly.

$$\begin{aligned} 65 &= \text{number of correct responses} \\ 110 &= \text{number of pupils answering the question} \\ \frac{65}{110} &= .59 \times 100 \\ &= 59 \quad \text{difficulty rating} \end{aligned}$$

In another school where the students had received the same instruction, 122 boys took the same test and 78 answered question 3 correctly Here the difficulty rating is 67.5, indicating that the test was easier for the students in this school

Questions with difficulty ratings above 90 or below 10 are not usually retained, for they are too easy or too hard. In the examples given above the ratings are useful. Question 3 is probably a good one. If the difficulty rating difference in favor of the second school happened on a number of questions, one might wonder if the instruction really was the same in both schools or one might look for some other influence in the school or community that might bear on the test answers.

Now for a further word on the discriminatory power of a test item. It has already been demonstrated that a "too easy" or "too hard" question does not help the teacher find out very much about the ability or knowledge of the students in the class. Carried to the extreme, if everybody fails or everybody passes a given test, then the test is hardly worth the time and effort spent to give it. It simply does not discriminate.

[11] Additional treatment of a mathematical nature will be given to the item of reliability in Chap. 18.

For a test question to have good discriminatory power, almost all the students answering the question correctly must rank above the students who answer incorrectly. In other words, if more generally poor students (on the total test) answer a particular question correctly than do generally good students, the question does not discriminate well. It has a negative discriminating power. If a teacher tested a class on a certain health lesson and a large number of the students who got As and Bs on the test failed one of the test questions, the teacher might want to look over the particular question rather carefully.

Multiple-choice items are often good discriminatory questions. By measuring how well they discriminate, it is possible to set up an *index of discrimination*. To illustrate, suppose that Oswego High School has just tested 100 students on archery knowledge. There were thirteen questions with multiple-choice items. For the purpose of studying question discrimination the following procedure is employed:

1. Select from the 100 examinations, the 25 papers with the highest scores and the 25 with lowest scores.[12] The high and low scores for the thirteen-question test, showing how each pupil marked the multiple-choice items, are shown in Figure 3-1.[13]

2. Tally all the scores on the tabulation worksheet in Figure 3-2, which has been prepared for the thirteen questions.[14] Note that each question provides for five to six multiple-choice items, a, b, c, d, or e. Tabulate the twenty-five highest cases first. For example, pupil 1 answered question 1 with c, question 2 with b, etc. These two tally marks are already placed in the appropriate cells of the table. Now tabulate the twenty five lowest cases using a colored pencil to prevent mixing high and low tally marks.

3. Make a mark in the "omit" column for any pupils that did not answer the question.

4. Compute the *difficulty rating* for each of the thirteen questions. Divide the total number of high and low tally marks in the correct cell by the total number of high and low cases. Record the rating in the appropriate place to the right.

5. If the difficulty rating is out of range place a cross (X) in the "revision" column.

6. To determine the *index of discrimination*, use the following steps:

a. Find the percentage of high scorers who answered the question correctly. In other words, what percentage of the twenty-five high-scoring pupils got the correct answer?

[12] Rather than handling all the test papers it is usually preferable to use approximately the 27 per cent with the highest scores and the 27 per cent with the lowest scores on the total test.

[13] In practice there is no need to list the test scores as shown. The twenty-five highest papers are merely piled together and tallied directly on the tabulation sheet.

[14] Permission to use the tabulation sheet has been obtained from M. Gladys Scott and Esther French. See reference to Scott and French at end of chapter.

Highest scores														Lowest scores													
Pupil	Question numbers													Pupil	Question numbers												
	1	2	3	4	5	6	7	8	9	10	11	12	13		1	2	3	4	5	6	7	8	9	10	11	12	13
1	c	b	d	c	d	e	a	b	d	b	a	e	e	1	c	b	d	c	d	e	a	b	d	b	c	a	e
2	c	e	a	b	d	c	c	c	d	a	b	c	d	2	c	a	d	c	e	d	a	b	d	b	c	a	e
3	c	a	d	c	e	d	a	b	d	b	c	a	e	3	d	d	a	a	b	b	c	c	e	e	a	b	c
4	d	a	c	c	e	d	b	a	c	b	a	e	e	4	c	a	a	e	d	c	b	a	d	c	c	a	e
5	b	b	d	a	a	a	a	d	b	a	c	a	a	5	a	b	c	c	d	e	a	c	b	e	e	a	e
6	d	e	a	c	b	d	b	a	d	b	c	d	e	6	d	d	a	c	b	c	d	c	e	a	a	e	a
7	c	a	b	d	c	b	a	e	d	b	c	a	e	7	a	d	e	a	d	c	d	a	c	c	b	b	c
8	a	c	c	e	d	b	a	d	c	b	c	a	c	8	e	c	e	b	a	e	d	e	d	d	e	e	e
9	c	b	d	e	a	c	c	d	d	e	c	a	e	9	d	d	d	c	e	d	c	a	a	e	d	d	c
10	a	b	c	d	e	d	a	b	a	b	e	e	d	10	c	e	c	d	a	b	c	a	d	e	e	a	b
11	c	a	d	c	e	a	d	e	a	b	c	a	e	11	a	a	d	e	d	c	b	b	a	e	d	d	c
12	a	a	d	d	c	c	e	e	d	b	c	a	e	12	b	d	b	d	e	a	b	c	c	d	c	e	a
13	c	a	d	b	d	d	e	a	a	d	e	a	...	13	b	a	b	c	...	b	a	c	e	d	b	c	b
14	a	e	c	c	d	d	e	a	c	c	c	c	c	14	d	b	c	d	d	e	a	d	b	c	a	d	e
15	c	a	d	c	e	d	a	b	a	b	b	c	a	15	e	c	a	a	b	a	e	d	d	c	a	b	a
16	e	a	d	c	e	d	a	b	d	b	c	a	e	16	c	d	e	e	c	e	b	e	d	a	a	c	e
17	c	a	d	c	c	d	a	b	d	b	c	c	a	17	a	d	a	e	d	a	c	b	e	b	b	d	a
18	c	a	d	c	e	d	a	b	d	b	c	a	e	18	a	d	c	a	a	d	e	e	c	b	b	a	b
19	c	a	d	c	e	d	a	b	d	b	c	a	e	19	e	a	c	b	c	d	e	a	d	a	d	d	c
20	c	a	d	c	e	d	a	b	d	b	c	a	e	20	e	e	d	d	a	b	c	a	d	e	d	a	b
21	e	a	d	c	e	d	a	d	d	e	a	a	a	21	d	b	d	e	d	c	b	d	a	e	e	d	c
22	c	d	d	a	a	d	c	c	a	b	b	b	e	22	d	b	e	c	e	a	b	a	c	d	e	e	a
23	c	a	d	c	e	d	a	b	d	b	c	a	e	23	c	d	a	b	...	b	a	c	e	d	c	e	b
24	c	d	e	c	e	d	a	b	d	b	c	a	e	24	e	d	a	b	...	b	a	c	e	d	d	a	b
25	c	e	b	d	c	a	a	d	b	b	c	c	d	25	e	d	c	e	...	d	a	e	c	e	e	c	d

Fig. 3-1. Scores on an Archery Knowledge Test.

b. Find the percentage of low scorers who answered the question correctly.

c. Translate the percentages answered correctly (for each question) into the index of discrimination and record on the table to the right. This is done by referring to Table A-7 (Flanagan's Table of Indices of Discrimina-

	a	b	c	d	e	Omit	Difficulty rating	Index of discrimination	Revision
1			/						
2		/							
3									
4									
5									
6									
7									
8									
9									
10									
11									
12									
13									

Fig 3-2 A Tabulation Worksheet (Note The correct answers are indicated by a heavy line around the proper cell, as in 1c, 2a, etc.)

tion) in Appendix A. High group proportions are on the horizontal axis; low group proportions are on the vertical axis. For example, if on a particular question the low group got 32 per cent right and the high group 94 per cent right, the index of discrimination would be 71.

7. Usually those questions with an index of discrimination of 20 or below

need revision. Place a cross (X) in the column marked "revision" where the index of discrimination is inadequate Figure 3-3 shows a completed table for the above problem. (The colored pencil marks are indicated by lighter-weight rule.) Note the number of questions in need of revision.

	a	b	c	d	e	Omit	Difficulty rating	Index of discrimination	Revision
1	//// 𝍸	/ //	𝍸 𝍸 𝍸 / 𝍸 /	// 𝍸 /	// 𝍸 /		44	64% 24% 41	
2	𝍸 𝍸 //// 𝍸	//// 𝍸 /	/ //	// 𝍸 𝍸	//// //		36	56 20 39	
3	// 𝍸 //	// //	//// 𝍸 /	𝍸 𝍸 𝍸 𝍸 /	/ ////		44	64 24 41	
4	// ////	/ 𝍸 /	𝍸 𝍸 𝍸 𝍸 //	𝍸 ///	/ 𝍸 /		44	60 28 33	
5	/// ////	// ///	//// //	//// 𝍸 ///	𝍸 𝍸 // ////	////	35	48 16 37	
6	/// ////	// 𝍸 /	/// 𝍸	𝍸 𝍸 𝍸 / 𝍸	/ 𝍸		42	64 20 46	
7	𝍸 𝍸 𝍸 / 𝍸 ///	// 𝍸 /	/// 𝍸	/ ///	/// ///		48	64 32 33	
8	//// 𝍸 //	𝍸 𝍸 / ////	// 𝍸 //	𝍸 ///	/// ////		32	44 16 33	
9	𝍸 ///	// //	/// 𝍸	𝍸 𝍸 𝍸 𝍸 ////	𝍸 /		12	0 24 −57	X — Too hard, does not discriminate
10	// ///	𝍸 𝍸 𝍸 //// ////	/ ////	/ 𝍸 /	// 𝍸 ///		46	76 16 60	
11	// 𝍸	/// ////	𝍸 𝍸 𝍸 / 𝍸	𝍸	//// 𝍸 /		42	64 20 46	
12	𝍸 𝍸 𝍸 𝍸 ///	/ //	𝍸 ////	/ 𝍸 //	/// ////		44	60 32 29	
13	/// 𝍸	𝍸 /	// 𝍸 /	/// /	𝍸 𝍸 𝍸 𝍸 //	/	45	60 28 33	

Fig 3-3 A Completed Tabulation Worksheet

Test Construction

Building an accurate test or instrument is not an easy task. This is especially true because few things in life can be proved; the best we can do is to present evidence The evidence collected by the test or measurement device must be of value Tests, therefore, are open to statistical and clinical appraisal; they must follow a proper pattern of construction.

No attempt will be made here to enter into a discussion of test construction as it applies to all the various types of measurement devices employed in

health and physical education. Only two types of devices will be considered: performance tests and written tests.

Constructing Performance Tests

Measuring the physical performance of boys and girls is one rewarding way of appraising health and physical education improvement in the school. Performance tests are widely used to measure elements of physical fitness: strength, muscular endurance, and cardiorespiratory endurance. They are used to note posture and body mechanics. They are used by the coach to check sports skills and motor coordination. The health teacher may use a bicycle skill test to make bicycle safety a meaningful subject. Performance tests are made up of tasks that require the manipulation of apparatus, mechanical equipment, human bodies, or other physical objects. Very often a performance test is so close to the actual element it is measuring that it is almost a reasonable substitute. The Celestial Navi-trainer, a device that presents the student navigator with a continuous view of the sky and stars as they would appear to him in a plane, is such a real test that for all practical purposes it is the same as actual flying conditions. As a performance test it rates very high. By the same token a basketball foul-shooting performance test might also rate high, especially if during a real game the foul-shooting ability of the individual player involved was very much the same as in the test situation.

Performance tests are usually put together in a battery or series of tests. Each separate item in the test battery must have value if the total test is to be effective. The steps in constructing a test battery are briefly as follows[15]:

Step 1. Select the test situations that seem to be in accord with performance goals.

If the purpose of the physical performance test is to measure capacities, status, or achievement, then some sample of performance must be set up. A basketball skill test, for example, might be put together by first listing a number of skills that in themselves would be a sample of performance. The considered judgment of a group of experienced persons would help to make the sample more representative of what it takes to demonstrate skill in basketball. A certain amount of trial and error may be necessary before the basketball skill test is complete with a number of proper test items. At this point some definite guides should be followed:

a. Each test item should measure a single reaction or simple, natural combination of skills but not a complex combination. When test items are simple, they are readily isolated and demonstrated and are thus more useful. It is easy to mix elements and draw erroneous conclusions. For example, a 50-yard swimming test

[15] The order established follows that outlined in American Association for Health, Physical Education, and Recreation, *Research Methods Applied to Health, Physical Education and Recreation*, National Education Association, Washington, 1959, p. 393.

can be set up to measure physical endurance, but because a boy cannot swim well he might be improperly rated as low in endurance despite a possible high level of general endurance and physical efficiency. In this case the test has not been made simple enough, an aquatic skill has been confused with physical capacity and endurance.

b. The performance test should measure one subject and his performance at a time.

c. The test item should measure abilities as nearly as possible in a life situation. A skill test should approximate a game situation.

d. Test scoring should be as objective as possible. The scoring system should yield a wide range of scores from high to low rather than a massing of scores at any part of the distribution.

Step 2. Determine the reliability of each test item. This can be done by correlating the scores on two successive administrations of the test on a large sample. The coefficient of reliability is achieved from the two performances. Tests low in reliability are discarded; others are remodeled.

Step 3. Determine the objectivity of each test item. The testing method used must yield a clear-cut value for each performance of the test item.

Step 4. Select the criterion for determining test validity. This can be done by comparing the test with a previously validated battery or by considering the rating of a group of judges.

Step 5 Compute the coefficients of validity by comparing performance scores on the test items with the scores on the criterion test.

Step 6 Compute the intercorrelations between all tests retained after Step 4. These separate sample situations or test items should strengthen one another in the total test but should not duplicate each other. If two test items correlated very highly, there would be little value in including both in the same battery, for it is possible that they both measure the same thing

Step 7. Put together into one test battery the best possible combination of test items from Step 6. This can be accomplished by employing the statistic of multiple correlation.

Step 8. Compute a formula for the test battery by which the separate test item scores can be combined

Step 9. Establish norms for the test battery. This means that a large number of subjects must be measured a group that belongs to the same population, with the same general likes and dislikes, opportunities and experiences. This is a representative group.

The nine steps outlined above represent radical simplification of a very exact and sometimes complicated procedure. A good part of this text could be devoted to each step and its development. This kind of test construction, especially as it involves performance tests, might well take several chapters in a book on advanced statistics and measurement. For the moment it is sufficient to say that good test instruments evolve over a period of time, a time filled with testing, retesting, and revision. Human nature and human behavior are much too complicated to be broken down into any easily defined sample. To isolate and limit human behavior items for purposes of

comparison individually or with other human performance is quite an undertaking. It is well worth the effort, however, for scientific analysis coupled with reasonable observations and judgments underlie the very basis of our present civilization. Appropriate evaluation is a continuous process in industry, business, medicine, and education. Any discipline that fails to use it because of the amount of work associated with it is doomed to failure, for to operate on a hit-or-miss basis is to invite chaos. Perhaps one reason that some health education and some physical education has appeared lame and impotent and has been labeled "frill courses" is that too little scientific evaluation has been brought to bear on the program as a whole. Where it has been properly applied, too little publicity has been given the results. If there were to be a tenth step in the above outline, therefore, it would be to make known in appropriate community circles the fruits of the measurement program and its meaning in the lives of growing boys and girls.

Constructing Written Tests

Written tests in health education have been used for a long time to find out what pupils know about a topic and occasionally to evaluate the more difficult items of attitudes and behavior. In physical education the use of short tests, quizzes, etc., is more recent and probably should be more widespread in regular class programs.

There has been an increasing use of objective tests in educational measurement.[16] These are sometimes referred to as guided-response procedures—a categorical term for true-false questions, multiple-choice items, matching questions, short-answer questions, and fill-in items. The term "objective tests," according to Bradfield and Moredock, has no official or even semiofficial standing as a technical phrase.[17] The word "objective" can refer only to a "single aspect of testing, that of scoring, while the construction and administration of such tests is as subjective a process as the construction and administration of any tests. Moreover, the phrase 'objective tests' connotes . . . that the procedures are free from human error, which, of course, they are not."[18] Although a subjective test is undesirable, criticisms of the objective examination are numerous. More than one author has pointed out that it encourages the wrong type of study habits (a bits-and-pieces approach); its items are frequently ambiguous; it facilitates cheating; and it seldom corresponds with the broad objectives of a course.[19]

It must be pointed out that a well-worded essay examination (free-response type) can be quite objective. It all depends on how the test

[16] For a discussion of the history of such tests, see Clay C. Ross, *Measurement in Today's Schools*, 3d ed., Englewood Cliffs, N.J., Prentice-Hall, Inc., 1954, chap. 2.

[17] James M. Bradfield and H. Stewart Moredock, *Measurement and Evaluation in Education*, New York, The Macmillan Company, 1957, p. 84.

[18] *Ibid.*, p. 84.

[19] F. L. Marcuse, "Objective and Subjective Examination," *School and Society*, 72:512–514, Aug. 26, 1950.

questions are worded. All proper questions are put together after the instructor has asked himself, "What decisions do I want the test results to help me make?" and "What data do I need in order to make these decisions?"

Some decisions useful to health and physical education instructors are the following[20]:

To determine a fair grade for each pupil based on over-all performance.

To be able to tell a pupil his pattern of strengths and weaknesses: has basic information been acquired? Can the student choose the right course of action? Can he support his choice by citing the appropriate principles?

To be able to identify pupils who have gained information but have not developed the ability to apply it.

To ascertain if there is a relatively high correlation between the ability to score well on information as well as application.

To be able to tell whether the instruction has been effective and where it is weak and in need of improvement.

With these decisions made, the test questions can be set up. Specific questions of either the free-response or choice variety can be put on paper. Their individual objectivity will be greatly improved if during the test-construction stage the following items are kept in mind:

1. The content of the test should range from very easy to difficult for the group being measured.
2. As a rule, the test should include more than one type of item.
3. It is usually desirable to include more items in preliminary draft of the test than will be needed in the final form.
4. The items should be so phrased that the content, rather than the form of the statement, will determine the answer.
5. All items of a particular type should be placed together in the test.
6. A regular sequence in the pattern of response should be avoided.
7. Ambiguous, trivial, and trick statements should be avoided.
8. Stick to one idea in a question, and try to stay away from partly true and partly false questions.
9. Avoid broad generalizations and the use of words such as never, always, none, and all. Such statements arouse doubts in the minds of intelligent pupils and are seldom true.
10. Later, when the test is actually tried out, every reasonable precaution should be taken to ensure normal conditions. The directions to the pupils should be as clear, complete, and concise as possible.
11. A simple scoring procedure should be adopted.

[20] Adapted from a report in Edward J. Furst, *Constructing Evaluation Instruments,* New York, Longmans, Green & Co., Inc., 1958, p. 194.

Free-response Questions

The supply or free-response question has considerable value in health education. It can test many outcomes; it minimizes the possibility of getting the right answer through guessing; it is relatively easy to prepare; and it reveals a degree of individuality. It includes the essay and short-answer type of questions.

Examples

Essay question: Describe the human heart and show its relationship to functional efficiency. Diagram the parts and relate its behavior to diet, exercise, and psychological stress. Be specific.

Short answer questions In a sentence or two express your understanding of each of the following terms:

conjunctivitis	impetigo
mitral valve	sodium fluoride
alcoholism	rheumatic fever
lumbar lordosis	vital statistics
wheat germ	somatotype

Choice Questions

The guided-response or choice questions are valuable because they do not depend upon skill in expression and handwriting. They force the student to select some answer, and they permit a relatively wide sampling of knowledge in a short time Once a key has been prepared, they encourage highly objective scoring.

Multiple Choice. The multiple-choice item consists of a stem or lead, together with a list of suggested alternatives, only one of which is correct or clearly the best. The alternatives are referred to as choices, options, or answer. The incorrect answers are referred to as foils or decoys.

This particular type of test question is extremely adaptable and lends itself to an unusually wide range of use. It is more reliable item by item than the true-false type and may provide more analytical data than true-false questions. A carefully formulated multiple-choice question with several alternatives can provide the teacher with a basis for appraising errors in thinking. The following points apply to the construction of multiple-choice test items:

1. Preface the questions with a short, clear set of directions.
2. State a single, definite problem in the lead.
3. Include as much of the item as possible in the lead.
4. Make the alternatives consistent with the lead.
5. Make the alternatives reasonably similar. The choices open to the student must be very much alike in order for the discriminatory power of the student to be measured. Too often there will be three or four choices, of

which all but one are obviously incorrect. In a five-choice question, at least three choices should be so close that only the student with real knowledge can select the most appropriate or best answer. The decoys therefore should be plausible and attractive.

Here are three questions used on a college examination in school health practice:

Example

Section I—Multiple Choice: Read each question carefully. Select the one item which *best* answers the question Put the number of the item selected in the space in front of the question.

______ 13. Proper school health permits children who have been ill to return to classes
1) after examination by the school physician.
2) after presenting a note from the family physician.
3) after being checked by the school or visiting nurse.
4) after first reporting to the classroom teacher
5) upon parent approval.

______ 18. Many studies, such as the Ohio Public Health Association School Health Survey, indicate that a large percentage of children never have a medical examination from one year to another The usual percentage given is:
1) 20 25%
2) 10-20%
3) 40-60%
4) 25-35%
5) 60-75%

______ 22. At what time of the year does the weight of school children increase more rapidly?
1) fall and early winter
2) spring and late summer
3) winter
4) summer
5) no set period

True-False. The true-false question should be conducted in the language of the pupil. It should be straightforward and as near to true or false as it can be made. It will be a better question when the statement has been prepared by the teacher rather than "lifted" directly from the page of the textbook or other instructional material. Also, sweeping generalizations should be avoided.

Example

Section II—True and False: The letters T and F have been placed before each statement given. Draw a circle around the letter T if the statement is *True* and around the letter F if the statement is *False* or *Partially False.*

T F 3. The more means the individual teacher has for observing and appraising behavior, the more likely she is to reach specific educational goals.

T F 12. The legal essence of the National School Lunch Program (Pub. Law 396, Congress 1946) is that the food supplied by the school must make a nutritional contribution to the meals.

T F 30. Accidents in the community are still the greatest threat to school children.

Sometimes in true-false questions it is a good idea to include information about scoring. The student may be asked to answer all questions, for to leave one blank will count against him. Here the test may be scored as right answers minus wrong ones (R — W), where the number of wrong answers (W) is subtracted from the number of correct responses (R). Some instructors, however, prefer to ask the students not to guess—not to answer a question when they do not know the answer To make certain that guessing is held to a minimum the pupil is penalized for getting the wrong answers. In this case a formula such as right answers (R) minus two times wrong answers (R — 2W) may be used Certainly this tends to discourage guessing.[21]

Affirm-Negate-Neutral. Somewhat related to the true-false test is the *affirm-negate-neutral* type of examination question. Here there is room for indecision, which in some cases is a proper way to answer the question.

Example

Section III—Affirm-Negate-Neutral Circle Yes, No, or U (Undecided).

1. Yes No U The trend now is to begin nutrition teaching in the early grades
2. Yes No U It is questionable whether sodium fluoride added to drinking water will reduce dental decay up to 60 per cent.
3. Yes No U Community health is everyone's business because the welfare of the majority is affected by the welfare of the minority.

Matching. Another kind of guided-response question is the matching variety. Matching items are easy to prepare and save space on the copy of the examination. Two columns are set up: one a "stimulus" column; the other a "response" column. These columns are located side by side with the stimulus on the left. There should be more response items than stimulus items. Directions should be clear as to which column is to be matched with which. If time does not permit for an oral discussion of this type of question, an example should be given on the examination sheet. These tests are very useful in measuring knowledge of sports skills, as well as in health examina-

[21] For an interesting coverage of the benefits of this method, see the article by Count Sussicran Etoxinod, "How to Checkmate Certain Vicious Consequences of True-False Tests," *Education,* 61:220–223, December, 1940.

tions where health occurrences are related to personalities, dates, elements, etc.[22]

Example

Section IV—Matching: Which of the sports terms listed in the left-hand column would you look first to find in the sports listed in the right-hand column? Consider each separately. Put the number of the sports term in the parentheses after each item.

1. Let ball	8. Golf	8 ( )
2. Par four	9. Hockey	9 ( )
3. Closed turn	10. Handball	10 ( )
4. Double dribble	11. Swimming	11 ( )
5. Icing the puck	12. Badminton	12 ( )
	13. Tennis	13 ( )
	14 Touchfootball	14 ( )
	15. Basketball	15 ( )

Completion, or Fill-in. The completion, or fill-in, item is one type of test question that requires the student to do something. In this case, instead of making a selection, the student is required to provide an answer. The degree of retention of a whole idea can be measured For example, a student's knowledge of an entire health concept can be tested by omitting the italicized words in the following sentence:

In the area of nutrition, the *Iowa Breakfast* Studies show that the *cereal* breakfast was nutritionally as favorable as was the *bacon and egg* breakfast.

Proper Measurement: A Challenge to the Profession

Professional workers in health and physical education often ask, "How technical should we become?" There is only one answer to this question, and that is to become just as technical as is necessary to ascertain whether physical, mental, social, and recreational fitness is being developed and maintained. The profession has come to the age really to move ahead and be of service to mankind. Human survival itself may depend more than most persons realize on the general acceptance of an evaluation program. What is accepted in any one community will brand unmistakably the status of health and physical education teachers in education and in the related fields of science and medicine.

All great professions have achieved status through measurement. One might even venture to state that the use of valid and precise measuring instruments is an absolute prerequisite for gaining recognition as a profession, whereas inaccurate and haphazard testing spells mediocrity, ineffectuality, and worse. Thus small and large schools alike should search for the best test available, begin in a small way to use it, improve upon it, and

[22] According to Clay C Ross, a matching test should contain at least five and not more than fifteen items, *op. cit.*, chap. 5.

evaluate the results. The challenge is to prove the value of the program. Failure to do so in the years ahead will result in labeling health teaching and physical education as "frill" courses, to be dropped from the already overcrowded school curriculum.

SELECTED REFERENCES

Bean, Kenneth L · *Construction of Educational and Personnel Tests*, New York, McGraw-Hill Book Company, Inc , 1953

Buros, Oscar K *The Third Mental Measurement Yearbook*, New Brunswick, N J., Rutgers University Press, 1949, pp 51–113, 475–480

Clarke, H Harrison *The Application of Measurement to Health and Physical Education*, Englewood Cliffs, N.J , Prentice-Hall, Inc , 1959, chap 2

Measurement and Evaluation Materials in Health, Physical Education and Recreation, National Research Council, American Association for Health, Physical Education, and Recreation, Washington, 1960

American Association for Health, Physical Education, and Recreation *Research Methods Applied to Health Physical Education and Recreation*, National Education Association, Washington, 1959

Rogers, Frederick R *Fundamental Administrative Measures in Physical Education*, Newton, Mass , The Pleiades Company, 1932, p 15

Ross, Clay C *Measurement in Today's Schools*, 3d ed , Englewood Cliffs, N J , Prentice-Hall, Inc , 1954, chap 1

Scott, M Gladys, and Esther French *Measurement and Evaluation in Physical Education*, Dubuque, Ia Wm C Brown Co , 1959

Troyer, Maurice E , and Robert C Pace *Evaluation in Teacher Education*, American Council on Education, Commission on Teacher Education, Washington, 1944, chaps 1 and 10.

CHAPTER 4

Evaluating the School Health Program

The successful school health program involves the combined efforts of many persons. It is a planned program involving administrators, teachers, and medical personnel It concerns the elementary school classroom teacher as well as the teacher of social studies, guidance, biology, home economics, health education, and physical education In larger schools it involves not only a physician and nurse but often a school dentist, school nutritionist, and psychiatrist actively engaged in promoting child health. These persons work in the three areas of health education: health services, healthful school environment, and health instruction

This cooperative effort to raise and maintain health status of school children is essentially educational rather than medical. This is another way of saying that the school does not practice medicine The school is concerned with health education experiences, the sum of which favorably influence habits, attitudes, and knowledge relating to health This text, therefore, is not concerned with the numerous medical and laboratory tests employed by medical personnel to find the cause of poor health; it is concerned with tests and measures that are primarily educational in nature and can be used by teachers and medical personnel alike.

Influential Elements in the School Health Program

Before one can properly appraise the effectiveness of the school health program it is necessary to break it down into a number of elements that influence it It is difficult to measure a total effect; it is not as hard to examine the individual parts Consider these influences:

Classroom Teacher. An alert classroom teacher can supplement the work of the physical education instructor and the school physician in a number of ways Very often, in the elementary grades, no other person sees the child quite as she does. If she uses her eyes, she can observe a variety of items related to the onset of poor health. The child with frequent colds and sore throats who it continually snuffing in the class may have any of a number of difficulties involving the upper respiratory area. Likewise, the child who squints to see or turns his head to hear simple questions may be in need of prompt medical aid. Numerous children appear to be well. The child who exhibits signs of emotional upset and is spotted by the teacher can be helped

at an early date. To the observing teacher, at any grade level, a particular response may represent a series of incidents leading to poor scholarship. The pupil, for example, who scores well on a test of learning aptitude yet produces work in class that is hardly better than average is the kind of pupil to watch.

Total School Environment. The school, to be at all effective, must be the kind of place a child wants to go. The classroom, cafeteria, playground, or gymnasium is made more inviting by the teacher. Such physical items as drinking fountains, hand-washing basins, shower facilities, ventilation, heating, temperature, lighting, furniture, and acoustical qualities are important school health considerations. In addition to these physical factors there are ever-present human factors—factors which are often more difficult to control. Not only are there individual pupil relationships to consider, but there is also the personal health status of the teacher. Personality maladjustments are responsible for failure in teaching more often than are physical disorders.

Administration. The value of the school administrator who is sympathetic to the goals of school health is great. He is most effective when he promotes in-service health education for medical personnel and classroom teachers, arranges for periodic nurse-teacher conferences and meetings of health councils, meets with parents to secure adequate treatment of children, assists in securing necessary health facilities, carries on studies relative to school absence and health problems, and provides positive leadership.

Physical Education Personnel. Total health is advanced by increasing the pupil's vigor, strength and endurance. Through a broad and varied program of physical activities organic growth is stimulated and fundamental physical skills are developed The physical education instructor, with his primary concern for motor activity and physical fitness, does much to enhance the total school health program. This is especially true when he works closely with the school physician and school nurse in adaptive or remedial programs.

Health Service Personnel. The duties and special influences of health service personnel are extensive The influence they have on the total school health program is far-reaching. In fact, it is these workers who make most of the health appraisal in many schools.

Coordinating Personnel. A number of persons in the school have a rather special function to perform. Their day to-day activities are coordinating. These persons include the health supervisor, health coordinator, school nutritionist, guidance specialist, and school custodian.

Health Council. Health counseling follows health appraisals. Health needs can often be served best by the combined attention of several interested persons. This group may consist of the school nurse, health supervisor or coordinator, the home-room teacher of the child under discussion, the parent, and possibly another person. Such working groups often appraise a situation and solve difficulties, which permits the school to give more than

"lip service" to meeting individual needs. There is in some communities another type of school health council, which is broader in scope than the committee already referred to. Such a council affords opportunities for widespread discussion and evaluation of health problems in a whole school system. Because it sometimes consists in teachers, administrators, parents, medical and dental personnel from the community, and representatives of health and welfare agencies, it is in a position to carry on a rather extensive evaluation project. More than one such council has surveyed the total school health program in an effort to measure its effectiveness.

Health Teaching. In the instructional program health habits, attitudes, and knowledge are taught By tests, observations, rating scales, surveys, etc., the conscientious teacher measures pupil progress. In addition, she periodically evaluates her own teaching methods by asking definite questions as to whether the health lesson was properly conceived and carried out.

Health Appraisal in General

There are many influences to consider in appraising the health program as a whole. The only tests and measures included in this chapter are those generally carried out under the immediate supervision of the health services department and the school administrator, such as routine medical examinations, numerous screening examinations, school surveys, health histories, and health records. All other catagories of health and physical education measures and techniques will be discussed in succeeding chapters.

As in all evaluation processes, the more scientific the appraisal, the sooner the purposes of the program may be reached Total health status of a student can be appraised in a number of ways Health histories, observations by teachers and nurses, medical and dental examinations, screening tests, posture tests, physical fitness measures, and psychological examinations are among the more common appraisal activities These appraisals serve the medical personnel, the classroom teachers, the physical instructors, and the pupils. They are also valuable to groups and individual persons concerned with health counseling.

A number of health appraisal programs are required by law. Such programs identify pupils in need of medical or dental treatment, those with nutritional problems, those who need psychiatric or child guidance help, those who are not growing properly, and those who may require modified programs of education, such as the crippled, the partially sighted, the partially deaf, and those with speech defects Often the appraisal program ferrets out pupils who need a more thorough examination than is usually provided at school, such as X-ray or laboratory examinations.

Medical Examinations

A good deal has been written about school medical examinations and how they can be administered and followed up efficiently. They must be compre-

hensive and painstaking if health status is to be accurately measured. In addition, they must be "sufficiently informative to guide school personnel in the proper counseling of the student, and sufficiently personalized to form a desirable educational experience."[1]

A thorough medical examination should be given either by the family physician or the school physician at periodic intervals in the life of the child. This is sometimes called the *routine medical examination.* A complete examination every two years is desirable, but it is often difficult to arrange because of the large numbers of school children and the frequent shortage of available physicians. The American Medical Association, in conjunction with the National Education Association, has suggested a minimum of four medical examinations: one at the entrance to first grade, one in the elementary intermediate school grades, one at the start of adolescence, and one before leaving school.[2]

An examination, to be worth the energy put forth, must allow the physician time enough to see the whole child: his behavior characteristics, nervous reflexes, speech defects, and nutritional status, as well as the more typical items. According to the Wisconsin Cooperative School Health Program, the minimum essentials of a good medical examination are the following[3]:

1. A check of the nose, mouth and throat for inflammatory discharges, obstructive growths or abnormalities.
2. Palpate for lymph node and thyroid enlargement.
3. Note deformities of the chest, spine and extremities. Note habitual posture and advise pupils at once if there are uneven shoulders, abnormal curvatures, protruding abdomen, or weak foot arches.
4. Using a stethoscope determine the condition of the heart and lungs. Tests should be made before exercise, then the pupil should exercise by doing "step-ups" or "push-ups" from the floor; then listen immediately for murmurs, heart rate, and check each lobe for rales; and after two minutes record the recovery pulse rate.
5. Test males for inguinal and umbilical hernia.
6. Examine the scalp, skin and between the toes for fungus growth.
7. Take blood pressure and collect samples of hemoglobin determination and urinalysis.

In addition to medical examinations given routinely during the school life of a child, there are times when it is necessary to give *special examinations.* These are examinations given to pupils before they compete in inter-

[1] National Conference for Cooperation in Health Education, *Suggested School Health Policies,* 3d ed., Chicago, American Medical Association, 1956, pp. 17–18.

[2] "School Health Services," *Report of the Joint Committee on Health Problems in Education of the National Education Association and the American Medical Association, School Health Services,* Washington, National Education Association–American Medical Association, 1953, pp. 40–43.

[3] Wisconsin Cooperative School Health Program, *School Health Examinations—A Guide for Physicians and School Authorities,* School Health Committee, State Medical Society of Wisconsin, Madison, Wis., 1951, p. 14.

scholastic activities and intramural sports activities. They may also be given to the pupil in adapted physical education, the physically handicapped, or the person who suffers sickness or injury, and as a result of information gained by special health screening procedures.

Health Histories and Cumulative Records

A review of individual pupil health histories and health records will frequently be most helpful to nurse-teachers, physical education instructors, guidance personnel, and health coordinators. In fact, school health records should be available for all teachers to look over. Meaningful information may be obtained, especially when the nurse is on hand to interpret screening tests and physical examination items.

The whole follow-up program depends on proper record keeping for concrete evidences of health improvement. The degree of personal illness, defects, and poor health habits; the extent to which remediable physical defects have been corrected; the improvement in growth factors; and the day-to-day anecdotal comments of teachers are typical points of information that contribute significantly to the health evaluation program.

A *health history* is a record of the pupil's illnesses over a period of time. It frequently includes a record of serious injuries and operations, an immunization record, and a place for noting special difficulties such as menstrual problems. History generally helps one understand the present better. Good teachers often ask "What makes Johnny tick?" In a number of schools today, classroom teachers help maintain health history records by anecdotally noting pupil health behavior.

Health histories and *cumulative records* go hand in hand. The results of all appraisals should be recorded in some kind of a cumulative pupil record folder. Such a health record might include the pupil's medical examinations, dental examinations, health history, observations by teachers, follow-up information on health improvement efforts, screening test results for posture, hearing, eyesight, nutritional level, and mental and emotional status, and the physician's report to the school administrator on significant medical findings. Haag points out that the record of the physician's report to the school is one of the most important parts of the health appraisal program.[4] One cannot act on the results of the evaluation unless the ordinary teacher can answer such questions as "Is the pupil capable of carrying a full program of school work?" and "Is special seating recommended?" The physical education teacher will also appreciate and be guided by information from the health record, which may be used to classify boys and girls for physical activity.[5]

In health evaluation, the record is a form of communication. It is or-

[4] Jessie Helen Haag, *School Health Program*, New York, Henry Holt and Company, Inc., 1958, p. 12.

[5] See classification chart, Chap. 14.

ganized, detailed, and condensed Accurate information from parents, teachers, and physicians is of value only when the program is set up to make use of the information Psychologists, attendance officers, speech therapists, and many other school specialists will make use of cumulative records that show a continuing picture Hinrichs, after a good many years of observation, points out that all too often school health records are inadequate to date This hinders school health surveys and creates large gaps of time between diagnosis, treatment, and follow-up[6] Because of this it is not surprising to find a number of states and communities periodically evaluating their records to see if they are still useful In Florida, for example, a continuous

ENTRIES by TEACHER, NURSE and PHYSICIAN

CODE—Worker Column T—Teacher N—Nurse Dr—Physician

TNC—Teacher Nurse Conference (If no symptoms noted entry after TNC should be NSN)

Date	Worker	Entries With Worker's Signature		Date	Worker	Entries With Worker's Signature

Fig 4-1 A Section of a Cumulative Health Record SOURCE Department of Health, Board of Education, New York City

effort has been made to make the health area of the Florida Cumulative Guidance Record (Grades 1 12) more effective by adding a section entitled "Significant Facts and Physicians' Notes"[7]

California Cumulative Record: Health Insert

In California a particularly useful health insert is used with the California Cumulative Record system[8] It is an integral part of the individual pupil's cumulative record and provides specific information concerning needed ad-

[6] Marie A Hinrichs, "Periodic Evaluation of Routines and Records as Milestones or Guideposts in the Progress of a School Health Program," *Journal of School Health*, 27 253–256, October, 1958

[7] For a copy of this record, contact the State Department of Education, Tallahassee, Fla (School Record MCH 304)

[8] This 1958 record can be obtained from Bureau of Health Education, Physical Education, and Recreation, State of California Department of Education, Sacramento 14, Calif.

justments in the school program. It is transferred from school to school, and it was designed so that a minimum of time is needed to keep it up to date. The health insert is divided into seven catagories, as follows:

1. Signs Observed by Teacher
2. Medical History
3. Immunization and Test Record
4. Growth Record and Screening Data
5. Dental Examinations
6. Medical Examination
7. Medical Recommendations for Needed Adjustments in School Program

New York State Cumulative Health Record

Over a period of years the New York State Education Department has revised and improved its Cumulative Health Record The first two parts of the six-part record pertain to the history and the annual health examination record. Parts 3, 4, 5, and 6 consider the treatment of defects, the causes of illness and absence, and a summary of observation.

Screening Tests

Measures employed by teachers and health services personnel alike to screen out or select those school children who appear to need further attention are referred to as *screening tests* Their sole purpose is to "weed out" persons who should be tested further by measures that require more time to administer, specialized medical personnel, complicated instruments, and facilities not normally found on the school premises

The most effective screening tests over the years have been tests of vision and hearing, tests for tuberculosis, and the appraisal of physical growth by the weighing and measuring of boys and girls. In recent years a number of school systems have attempted to screen children for nutritional status, dental defects, and level of mental health Most of the really practical screening tests can be administered by teachers and nurses. In many secondary schools, and more and more elementary schools, physical education personnel are screening for postural defects, evidence of low physical fitness, and inadequate neuromuscular skill

Because of the shortage of medical personnel in certain regions of the United States the need for screening examinations is magnified considerably. Research reveals that school screening examinations of high quality ensure referrals and follow-ups.[9]

Vision Testing

To identify pupils with vision problems, school programs rely primarily on two procedures: (1) nurse and teacher observations and (2) screening tests

[9] Report of the Sixth National Conference on Physicians and Schools, *Journal of Health, Physical Education and Recreation*, 29:22–25, February, 1958.

for visual accommodations. Each of these is needed; in fact, each supplements the other.

Some 80 per cent of the work a child does in elementary school requires visual acuity within arm's reach, and three-fourths of all study activities in most schools require some reading ability The early detection of visual difficulties by alert teachers is most important. It is a continuing process from day to day; it can be on occasion more effective than the periodic screening tests for vision. A list of deviations in appearance and behavior, which may be used as a kind of check list, is shown on page 54.[10]

Deciding upon a screening test for visual acuity is the responsibility of several persons—perhaps a local committee composed of ophthalmologists, optometrists, school medical advisers, nurses, and educators.

When a pupil is old enough to cooperate in a testing situation, he is ready to be measured. Thus, even a three-year-old child can respond to the symbols on a Snellen E chart Of course, proper advance preparation helps improve response It should be added, however, that no matter how well the pupil is prepared or how well the test is given, no screening examination is completely satisfactory Screening simply lets the school personnel know that there is some visual difficulty; further examination will be necessary by a qualified eye specialist

The *Snellen test*, used for a long time in the public schools to test visual acuity, is quite satisfactory as a crude screening instrument when more valid measures are not available[11] Certainly it is not the most discriminating measurement of vision, for it does not test eye coordination, astigmatism, balance, etc. It is useful for noting near-to-far vision It will tell the classroom teacher whether or not a certain pupil can see the blackboard or wall chart from either the front or the back row of seats The large Snellen test chart requires reading test objects (letters, numbers, or symbols) on a chart from a 20-foot distance The E chart is available for children and illiterates, and the mixed-letter chart is for children and adults above the third-grade level The largest letters should be read by the normal eye at 200 feet. Because testing is done at 20 feet, this line of large print is designated the 20/200 line. The normal eye should read the 20/20 line at 20 feet; deviation less than normal is recorded as 20/200, 20/100, 20/70, 20/60, 20/40, 20/30, and any deviation better than normal is recorded as 20/15 or 20/10. In all probability children below the third grade who cannot read the 20/40 line with each eye should be referred for examination by a specialist.

Although the Snellen test can be administered by a person without medical or nursing training, proper attention must be given to the testing pro-

[10] Adapted from the report, "A Guide for Eye Inspection and Testing Visual Acuity," National Society for the Prevention of Blindness, 1790 Broadway, New York 19, N.Y.

[11] The necessary testing equipment can be obtained from the National Society for the Prevention of Blindness, 1790 Broadway, New York 19, N.Y., or the American Medical Association, 535 N. Dearborn St., Chicago, Ill.

Eyes	Always	Usually	Seldom
Complaints of dizziness, headaches and nausea			
Crusts on lids among lashes			
Red rims on lids			
Watery eyes; styes			
Swollen eyelids			
Reading difficulties: Holds book far away from face when reading			
Holds body tense or thrusts head forward at distant objects			
Holds face close to page when reading			
Inattentive in wall, chart, blackboard, or map lesson			
Reads but brief periods without stopping			
Screws up face when reading or looking at distant objects			
Shuts or covers one eye when reading			
Tilts head to one side when reading			
Tends to look cross-eyed when reading			
Tends to lose place on page			
Tends to make frequent changes in distance at which book is held			
Confusion in reading and spelling o and a; e and c; h, n, and r; f and t			
Apparently guesses from quick recognition of parts of the word in easy reading material			
Reversal tendencies in reading			
Rubs eyes frequently			
Poor alignment in writing			
Attempts to brush away a blur			
Irritable over work, or some emotional display			
Strabismus, or squint			

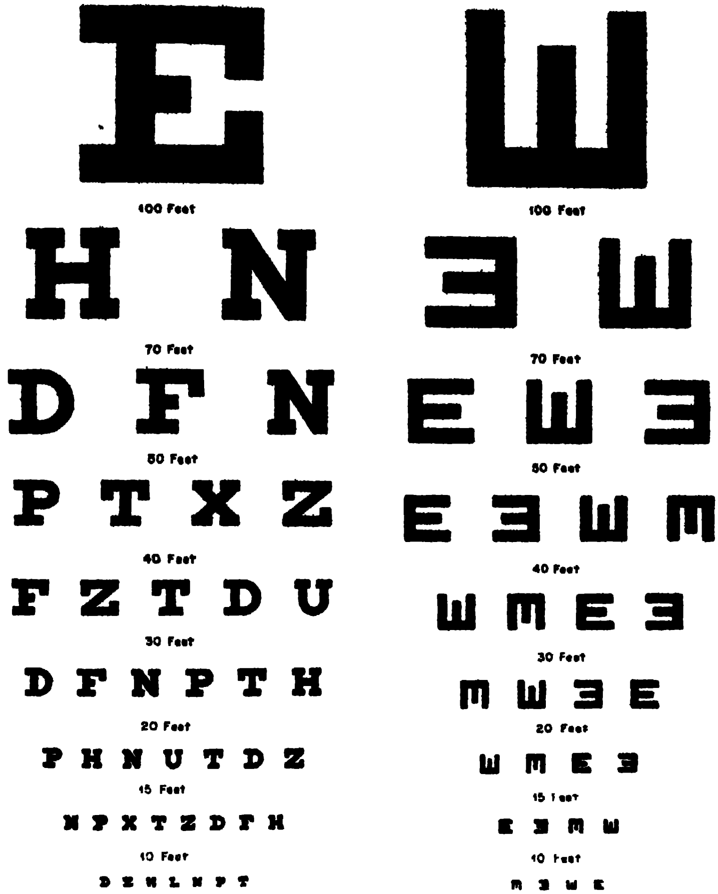


Fig. 4-2. The Snellen Charts. SOURCE National Society for the Prevention of Blindness, 1790 Broadway, New York 19, N.Y.

cedure and equipment. The testing kit and instructions can be obtained from the American Medical Association or the National Society for the Prevention of Blindness. Some of the more important items to keep in mind to ensure valid results are the following[12]:

Hang the chart so that the 20-foot line of letters is at the level of the child's eyes.

The heels of the child should touch the line marked on the floor 20 feet away.

Children should be at ease and encouraged to do their best.

Both eyes should be open during the test, the eye not being tested covered with a small card or folded paper resting obliquely across the nose.

[12] Adapted from "School Health Services," *Report of Joint Committee on Health Problems in Education of the National Education Association and the American Medical Association,* Washington, National Education Association–American Medical Association, 1953, pp. 76–77.

Test the child who wears glasses first with them and then without them.
Test the right eye first, then the left, and then both eyes together.
Begin with the 30-foot line and follow with the 20-foot line. If he fails the 30-foot line, start with the 20-foot line.
Keep unused parts of the chart covered, using cards with circular cutouts to expose single letters. This prevents memorization.
Move promptly and rhythmically from one symbol to another at a speed with which the child can keep pace.
Consider a line read satisfactorily if three out of four symbols are read correctly.
Record results immediately in fraction form, the numerator representing distance from the chart (20 feet) and denominator representing the lowest line read correctly.
Look for eyestrain; such signs as excessive blinking, frowning, scowling, tilting the head, and watering of the eyes.
Correlate Snellen test results with classroom observations.

The *Keystone View Company Telebinocular Test* is recommended when a more complete screening examination of eye functions is desired.[13] In recent years it has been used not only for routine screening examinations but to measure driver vision in driver-training programs. The Telebinocular is a kind of stereoscope with stand, slide holder, and light. It is used with the Betts Visual Sensation and Perception Tests, which have been set up in slide (card) form.[14] When viewing the cards in the Telebinocular the eye sees only its half of the slide; this permits the eyes to be tested together or separately. In this test it is possible to screen the eyes for such functions as vertical balance, lateral balance, fusion, visual acuity, usable vision of both eyes together and separately, depth perception (stereopsis), color perception, and peripheral vision.

There is varied opinion among ophthalmologists on the value of binocular devices. Some like them; others consider them unnecessary if the Snellen tests are used in conjunction with parent-teacher observations. It is apparent that a good observation is a fair form of evaluation.

The *Massachusetts Vision Test* was developed by the Massachusetts State Department of Health as a more efficient device for the detection of faulty vision among school children.[15] It works well at all school ages. It measures the degree of low visual acuity, notes the lack of binocular coordination (that is, where there is difficulty using both eyes), and detects conditions causing eyestrain. It uses different lenses. A convex lense is placed in a testing frame, and the pupil is asked to read the 20-foot line on a Snellen chart.

[13] Keystone View Company, Meadville, Pa.
[14] Emmette A Betts, *The Prevention and Correction of Reading Difficulties*, Evanston, Ill., Row, Peterson & Company, 1936, p. 161.
[15] Available from Welch Allyn, Inc., Auburn, N.Y.

If he can do this, he fails the test, for with the lens in front of his eyes he should not be able to see it clearly. Reading the 20-foot line is related to the known strength of the lens. A Maddox rod and two pair of spectacles are used in the second test. One pair of spectacles produce a horizontal streak, the other a vertical streak. The pupil views a picture of a house drawn on a chart on which a light is placed in a small opening corresponding to the window of the house. The house is built to scale so, if a light streak is

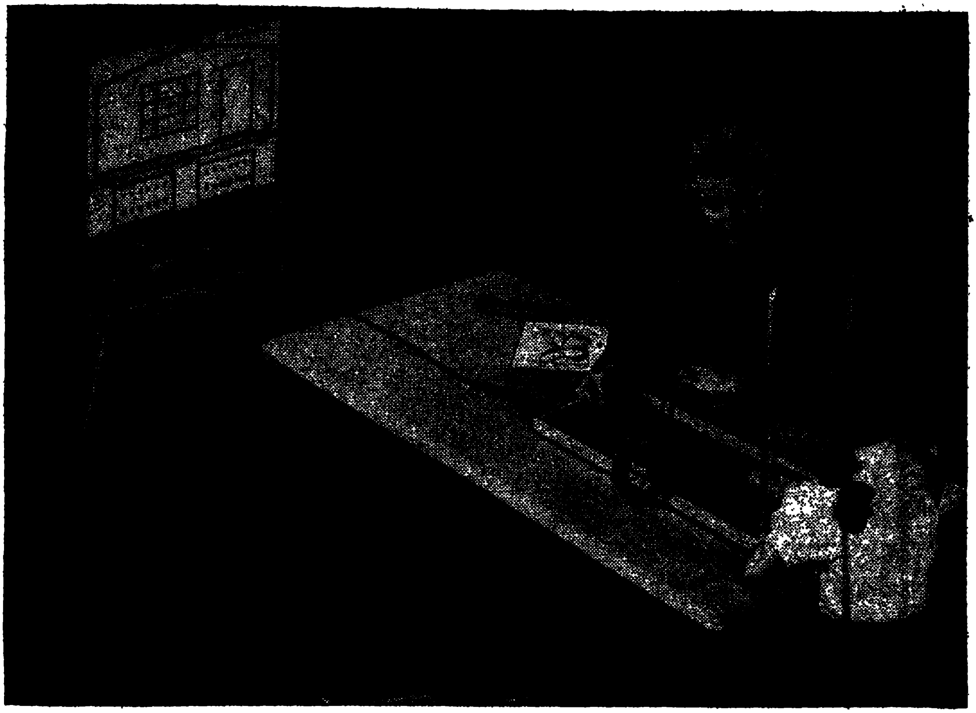


Fig. 4-3. The American Optical Vision Screening Test. A portable carrying case, which unfolds into the floor and desk units, is included. SOURCE: American Optical Company, Instrument Division, Southbridge, Mass.

seen outside a specified area, a significant degree of muscular imbalance is indicated.

The *American Optical Vision Screening Test* was designed as an improved version of the Massachusetts Vision Test.[16] It may be given by a competent lay person or nurse-technician. It takes less time to administer than the Massachusetts Vision Test. When a first failure is encountered in any of the seven tests (lateral phoria, vertical phoria, binocular acuity, left acuity, right acuity, depth perception, superimposition, and abduction), the screening is considered as complete. Visual acuity is checked under several conditions at the standard 20-foot distance. Factors contributing to

[16] Available from American Optical Company, Instrument Division, Southbridge, Mass.

eyestrain are noted at both the 20-foot test distance and at the normal 16-inch reading distance. In the American Optical device there are no test spectacles to put on the pupil; instead, the occluders and optical test elements for both right and left eyes are mounted in a self-supported headpiece. The examiner simply dials each successive test into position before the student's eyes. This is a well-engineered test.

The *Ortho-Rater*, used extensively in the accident-prevention clinics of the New Jersey Division of Motor Vehicles, involves several pieces of electrical equipment.[17] It is particularly useful in driver education because it sets up practical situations. For example, in the reaction-time test the subject is required to move his foot from the accelerator to the brake. Stopping distances at given speeds are determined The depth perception test weighs errors of driving judgment, such as cutting in too sharply or following too closely. The seven tests in the Ortho-Rater battery are visual acuity, field of vision, depth perception (day and night), simple reaction time, traffic color recognition, night vision, and glare recovery.

Validity of Vision Screening Tests

In the St Louis, Missouri, schools an interesting piece of research was carried out to ascertain which of the several vision screening tests were the most valid in detecting visual difficulties[18] Test findings were related to the clinical judgment of three ophthalmologists Among the findings were the following points of interest to school personnel·

a. There was a low degree of agreement between referrals by ophthalmologists and those by any of the screening procedures used

b. The study did not indicate that one screening procedure was better than any other procedure.

c. Visual screening for beginning readers appeared to be worthwhile

d. Individual school systems must decide on the screening device best suited for their programs

e. Screening procedures are not positive proof that a pupil has a defect.

It is worth noting that the objectivity and reliability of these tests was satisfactory. As consistent results were obtained by different testers, it seems safe to suggest that elaborate training of testers is not necessary[19] The same kind of results were obtained in Maryland when a number of persons of varying abilities administered both vision and hearing tests[20] This point alone should encourage more extensive use of screening tests

[17] Available from Bausch and Lomb Company, Rochester, N Y

[18] Marion M Crane, "Study of Procedures Used for Screening Elementary School Children for Visual Defects," *American Journal of Public Health,* 42 1430–1439, November, 1952

[19] Marion M Crane, Richard Scobee, Franklin W. Foote, and Earl L. Green, "Study of Procedures Used for Screening Elementary School Children for Visual Defects," *Journal of School Health,* 23·1, January, 1953; 23·44, February, 1953.

[20] Mary A. Thompson, "Vision and Hearing Screening Program in Prince George's County, Maryland," *American Journal of Public Health,* 47:200, February, 1957.

In an effort to keep vision screening as simple and efficient as possible, a survey by the Nebraska Optometrists Association recommended that only two tests be given: (1) an annual acuity test for all grades, with referral levels of 20/40 for grades kindergarten, one, and two; 20/30 for grades three and four; and 20/20 for grades five through twelve; and (2) hyperopia tests for selected cases referred by teachers on the basis of reading difficulties.[21]

Certainly visual acuity testing has proved worthwhile in American schools. If it can be limited to a simple procedure, more schools will employ visual testing. An alternative procedure is to bring the medical specialist or optometrist to the school to do the screening, a modified clinical technique.

Hearing Testing. Much of what has been written about tests of vision applies as well to tests of hearing. Early and repeated examinations are important in the program. Accurate screening is necessary especially when one considers that an average of two children in every classroom have a hearing problem, that many cases of incorrect speech supposedly due to poor mentality are actually caused by poor hearing, and that there are needless grade repetitions among hard-of-hearing children.

As in vision appraisal the alert and sympathetic teacher is more useful than several testing instruments. She may observe and report such items as discharge or odor from the ears, earache, failure to answer, misunderstanding, habit of saying "What?," turning head to one side when spoken to, facial expressions indicating a lack of awareness of what is going on, faulty pronunciation of common words, and speech with an unusual sounding voice.[22]

The *watch-tick test* and the *whisper test*, which are based on the ability of the pupil to hear the tick of a watch or the whisper of the teacher's voice, are much too crude for present-day use. Group and individual audiometers are more valid. It should be noted that as new and improved audiometers are put on the market, an approved list of audiometers is revised by the Council on Physical Medicine and Rehabilitation of the American Medical Association.[23]

The *group audiometer*, or phoaudiometer, is designed to test ability to hear numbers spoken in progressively decreasing intensity. It is somewhat like a recording machine, with up to forty sets of earphones permitting as many as forty pupils to be tested simultaneously. A record is played in which a series of numbers are called by first a man and then a woman. As the voice continues, the numbers are given in a lower and lower tone. The pupil, listening through his earphones, writes down the numbers as he hears them,

[21] John C. Thompson, "Selecting a Vision Screening Instrument for Lincoln, Nebraska Public Schools," *Journal of School Health*, 28:250, October, 1958.
[22] Carl E. Willgoose, *Health Education in the Elementary School*, Philadelphia, W. B. Saunders Company, 1959, p. 119.
[23] Current lists are available from the American Medical Association, 535 N. Dearborn St., Chicago 10, Ill.

entering them in the appropriate column on a test form. After each ear has been tested separately, the test form is turned over to the tester for scoring. Anything less than a loss of 9 decibels is considered normal.[24] Above this level the pupil should have an examination by a specialist.

The *pure-tone audiometer* is superior to any other hearing screening device today, particularly if used with the sweep-check method. There is a set of earphones for the subject. Sound intensity is regulated by a knob up to maximum intensity. Different pitches in tone are given (frequency range) by turning another knob. This threshold testing gives a fair indication of a student's hearing loss, provided extraneous noise and other distracting elements are eliminated while the testing is done. Since the standard pure-tone audiometer scale is logarithmic, each step of 10 decibels represents ten times the previous step in decibel loss. If the loudness of a sound has to be stepped up to 10 decibels in order to make audible a sound that normally would be heard at 1 decibel, the amount of hearing loss is expressed as 10 decibels. A loss of 15 decibels or more in two or more tones indicates a significant hearing impairment.

This test may be given to 100 pupils a day without difficulty. It is most accurate on an individual basis, and it can be used with children of all grades. Several companies put out a model for group testing in which multiple headphones are provided.

In some schools a *pure-tone threshold acuity test* is given to pupils who fail a pure-tone audiometer test.[25] The threshold acuity test discovers the given intensities needed for a particular pupil to hear tones of varied pitch.

The need for hearing testing is greater than most persons realize. Hearing loss in children worsens as the years pass. Many children with partial hearing loss are as much as two years behind their classmates.[26] Some pupils fail the 4,000 to 8,000 high range; others fail the range below 500; but it is especially important for them to hear in the middle range. Across the country 5 per cent of all children tested fail screening tests in hearing. Of this 5 per cent, about 40 per cent have a true hearing loss. Unfortunately, only 15 to 20 per cent of children with a known loss are receiving medical and educational help.[27] There is little doubt that hearing problems can be greatly re-

[24] The decibel is a unit of measure applied to hearing It is equal to 1 bel, the unit of energy for the transmission of sound by electricity

[25] Texas State Department of Health and Texas Education Agency, *A Hearing Conservation Program for the Children of Texas*, Austin, Tex., Texas State Department of Health, 1957, p. 12–18

[26] Samuel M. Wishik et al., *Audiometric Testing of Hearing of School Children:* II. An Eight Year Longitudinal Study of Hearing Status of 1,726 Children, Including Prognostic Implications and Relationship to Academic Retardation, presented at the annual meeting of the American Public Health Association, Atlantic City, N.J., Nov. 14, 1956.

[27] From an extensive report by Dr. C. Adele Brown, of the Committee on Hearing Conservation, School Health Section, American Public Health Association, presented at the annual meeting in St. Louis, October, 1958.

duced if screening is done in the lower grades. As it is, hearing referrals and follow-up practices are very poor and not uniform throughout the country.[28]

Screening for Physical Growth

There is a close relationship between pupil health status, personal growth characteristics, and academic achievement. Learning, in its many forms, depends upon sound health. Very often a simple sign of retardation in growth, discovered by nurses or teachers, is significant. Yet the growth period of boys and girls cannot be nicely divided and subdivided into set periods There is too much individual variation in pupil size, shape, and form due to such factors as heredity, constitutional endowment (body build), and physiological maturation A number of screening devices to evaluate over-all physical growth have been developed for school use

Age-Height-Weight Measurements. There is value in using age-height-weight tables, provided one keeps in mind that for any given sex and chronological age there are wide differences in individual pupil maturation and type of body build Probably the child whose measurements fall within 10 per cent of what the norm table calls for is within the normal range. Special tables that consider height and weight in terms of light frame, medium frame, and heavy frame are much more revealing than the single average of any one age and sex[29]

The child that fails to gain in height and weight in a four-month period frequently has some degree of ill-health This may be due to a physical defect, an organic drain, a prolonged emotional problem, or perhaps a number of unhygienic health habits Turner points out that height and weight should be screened every month because practically all elementary school children (98.4 per cent of 4,200 observed) fail to show a gain on at least one monthly weighing during the school year[30]

Weighing at the school should be done at approximately the same time of day each month Scales and measuring rods should be tested frequently to be sure they are accurate Children remove hats, coats, jackets, sweaters, and shoes The weight is recorded in pounds and half pounds Height is recorded in inches and quarter inches For practical purposes an individual record card for pupils in the lower grades should be kept by the classroom teacher In the upper grades the records are usually maintained in the health service office.[31]

[28] *Ibid.*, p 8

[29] Tables are available from the Metropolitan Life Insurance Company, 1 Madison Ave., New York, N Y

[30] Claire E Turner, C Morley Sellery, and Sara Louise Smith, *School Health and Health Education*, St Louis, The C V Mosby Company, 1957, p 102

[31] A useful individual weight-height record form for recording height and weight each month of the school year can be obtained from the American Medical Association, 535 N. Dearborn St., Chicago 10, Ill.

Over a number of years the *Meredith Physical Growth Record*, individual height-weight-age graphs for boys and girls developed by Howard V. Meredith, have proved especially useful in appraising pupil growth [38] There is a record for each sex, which is divided into zones. For height there are five zones: tall, moderately tall, average, moderately short, and short. For weight the zones are heavy, moderately heavy, average, moderately light, and light. The graphs cover ages four to eighteen.

[38] For a complete description, see Howard V Meredith, "A Physical Growth Record for Use in Elementary and High Schools," *American Journal of Public Health*, 39·878–885, July, 1949

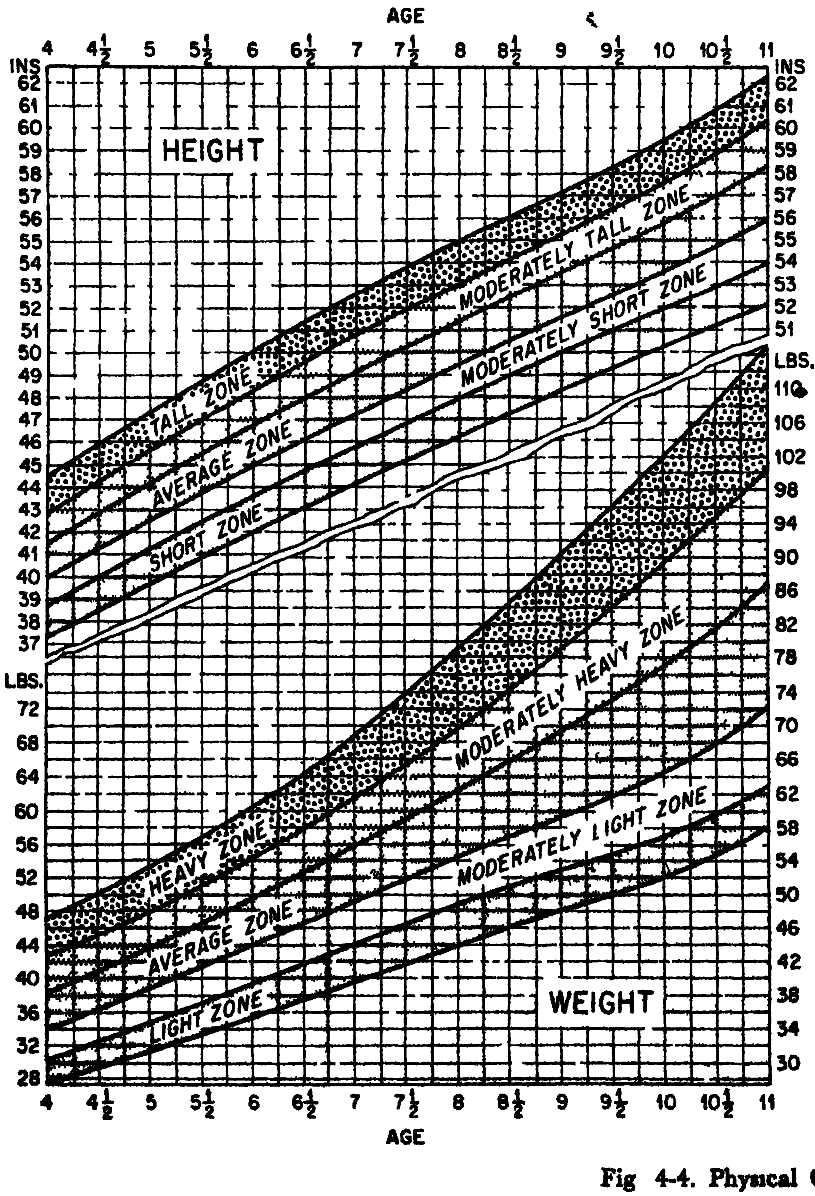
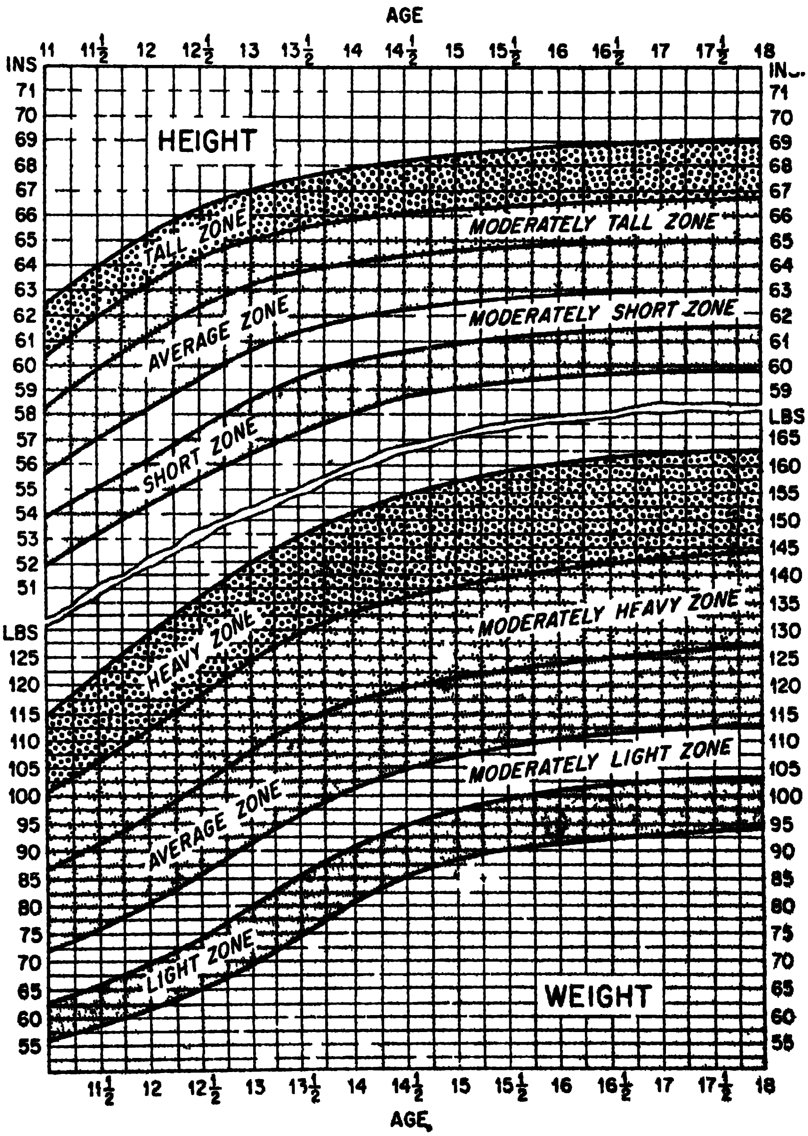


Fig 4-4. Physical Growth

The child's measurements are plotted on the graph for any one period. Successive measurements permit the observer to note whenever the points do not fall in *like zones* (that is, tall and heavy, short and light). In such instances the pupil is referred to the physician for examination. It may be that the pupil is perfectly healthy but of a particular body build. On the other hand, he may have some infection, need an improved diet, or require changes in daily living habits.

The *Wetzel Grid* is a screening procedure based on the principle that normal growth proceeds along a channel characteristic of the body build of



Record for Boys.

the individual child. The charting on this grid is more complicated than on the Meredith graphs, but it is also more revealing of pupil health status.

In a number of public schools the Wetzel Grid has proved useful in the school health examination when used under the supervision of the school physician.[11] It is sometimes listed as an instrument that measures nutritional

[11] Norman C Wetzel, "The Simultaneous Screening and Assessment of School Children," *Journal of Health and Physical Education,* 13 576, December, 1942, see also *The Treatment of Growth Failure in Children,* Cleveland, Ohio, National Education Association Service, Inc , 1948

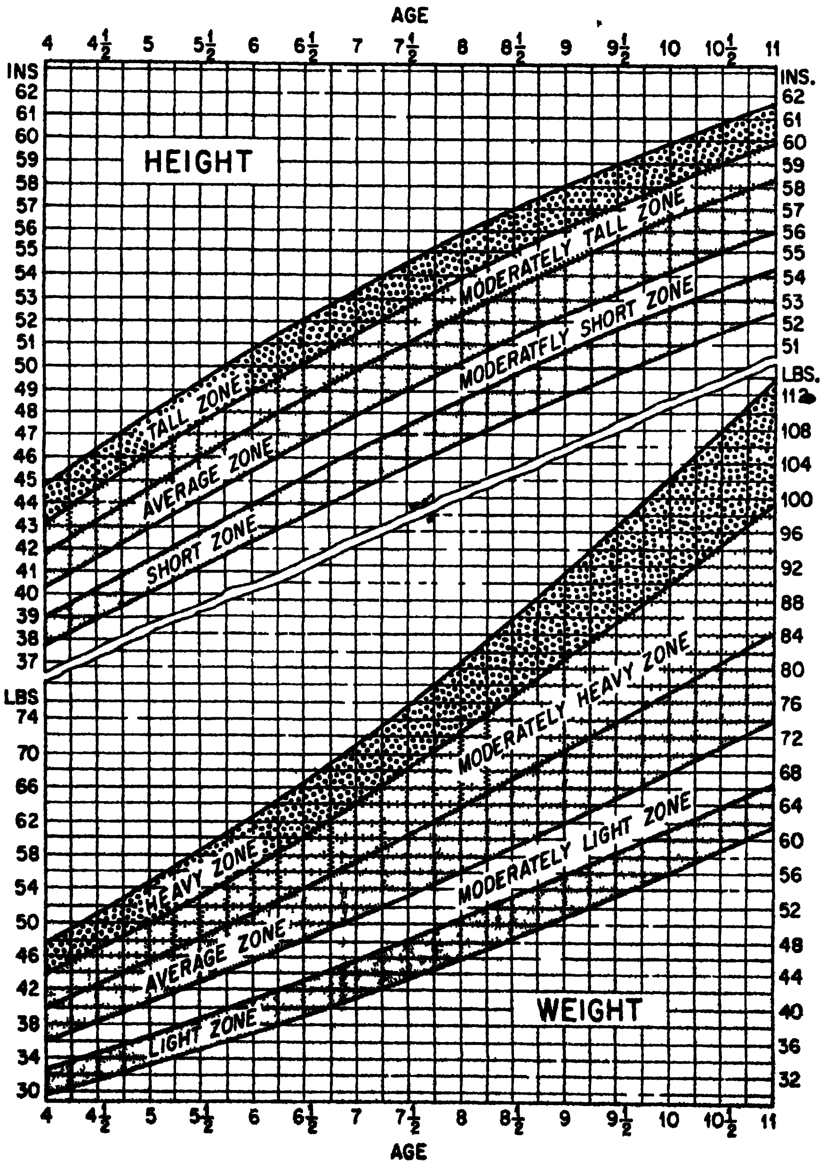
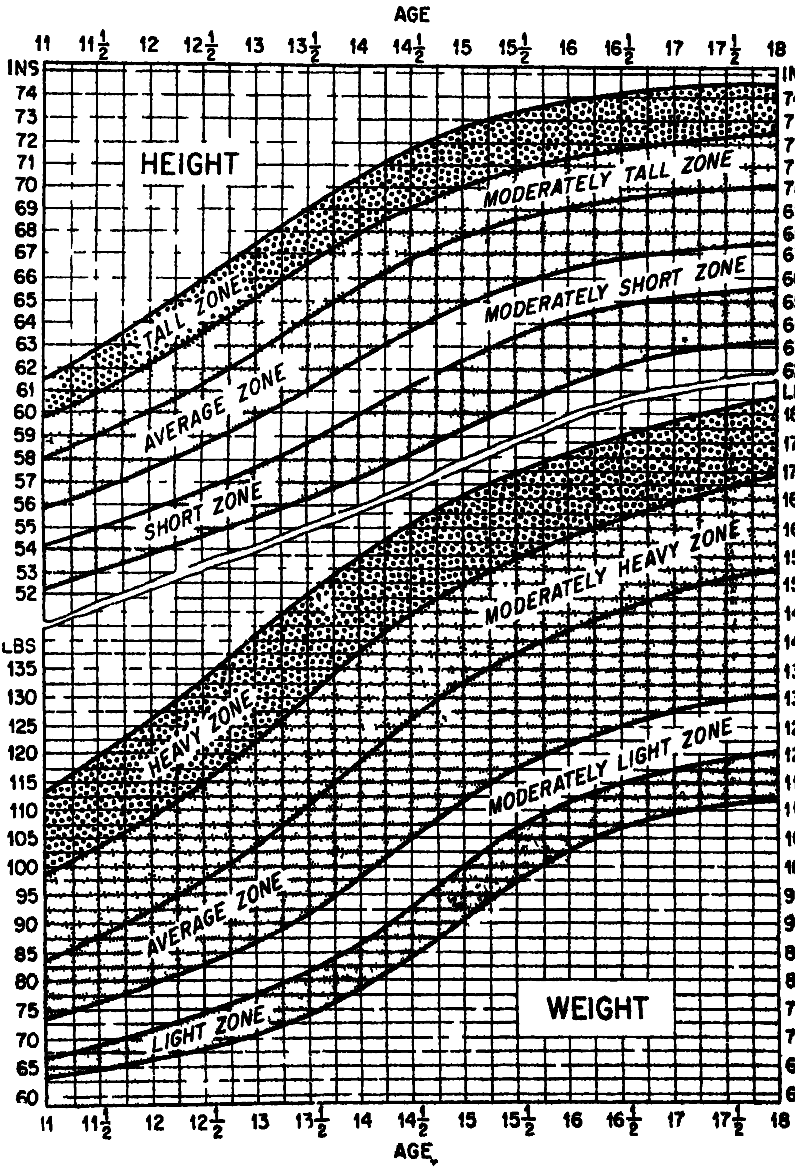


Fig 4-5 Physical Growth

status. As a nutritional measurement it detects abnormal trends toward obesity or underweight. Pugh, in an extensive study of children at the laboratory school at The Ohio State University found that the grid had definite value as part of the child's school health record and was especially useful as a counseling aid with students, parents, and teachers.[34]

Wetzel worked out a gridlike chart that describes seven normal hereditary types of body shape, from obese to slim, and mapped what he called

[34] Molly C Pugh, "Charting Growth with the Wetzel Grid," *Journal of Health, Physical Education and Recreation,* 25 48, March, 1954



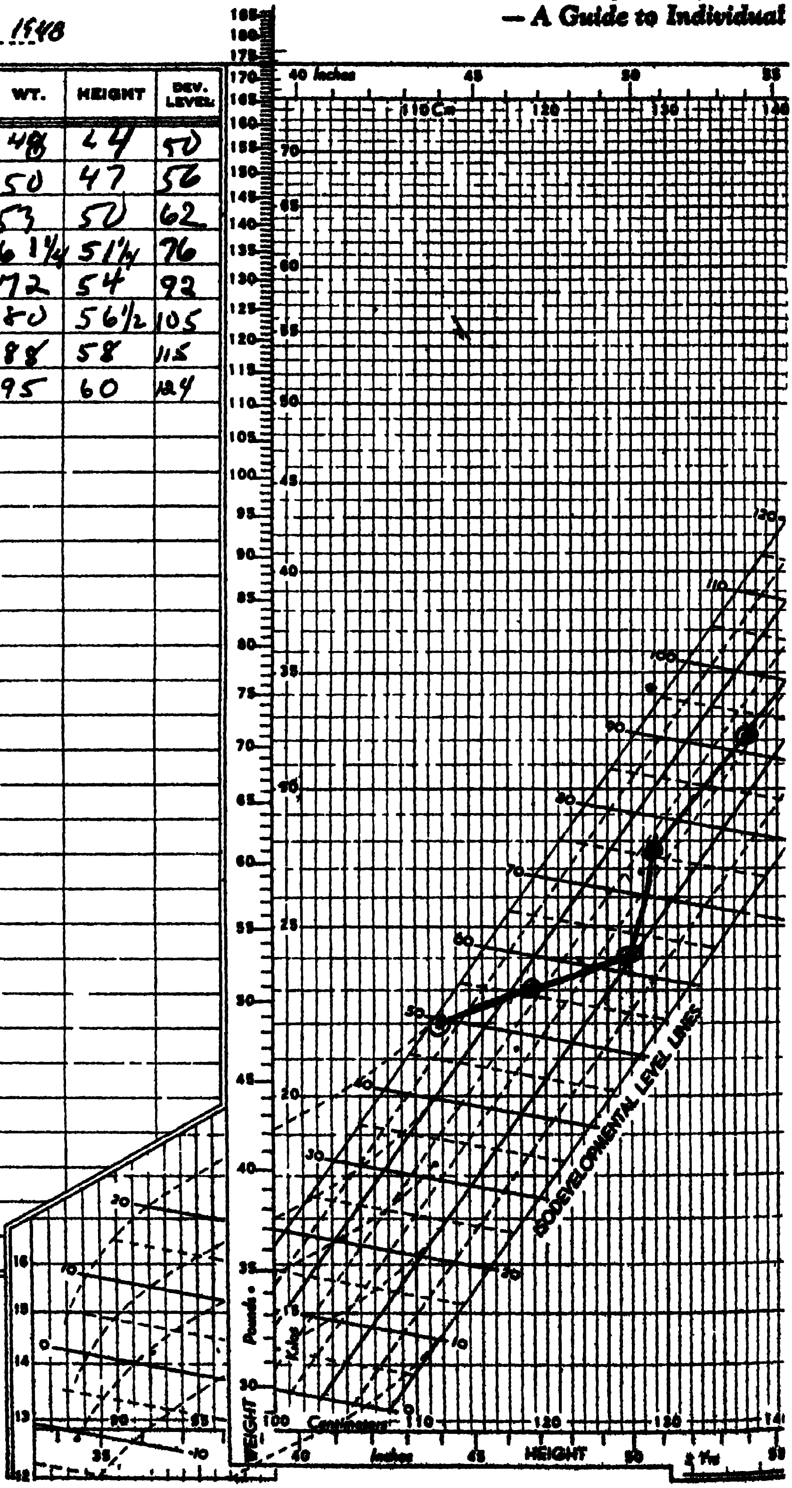
Record for Girls.

GRID for Evalu
in Terms of PHYSIQUE (Body Build), I
— A Guide to Individual

Name *Tommy Willgoose*

DATE OF BIRTH *26 Feb. 1948*

DATE	AGE	WT.	HEIGHT	DEV. LEVEL
[illegible]-53	5	48	44	50
[illegible]-54	6	50	47	56
[illegible]-55	7	53	50	62
[illegible] 56	8	61¼	51¼	76
June 57	9+	72	54	92
June 58	10+	80	56½	105
June 59	11+	88	58	115
June 60	12+	95	60	124



TOLERANCE LIMITS

DIRECTION—½ channel per 10 levels of progress

SPEED: 2 to 3 levels from own auxodrome in any year

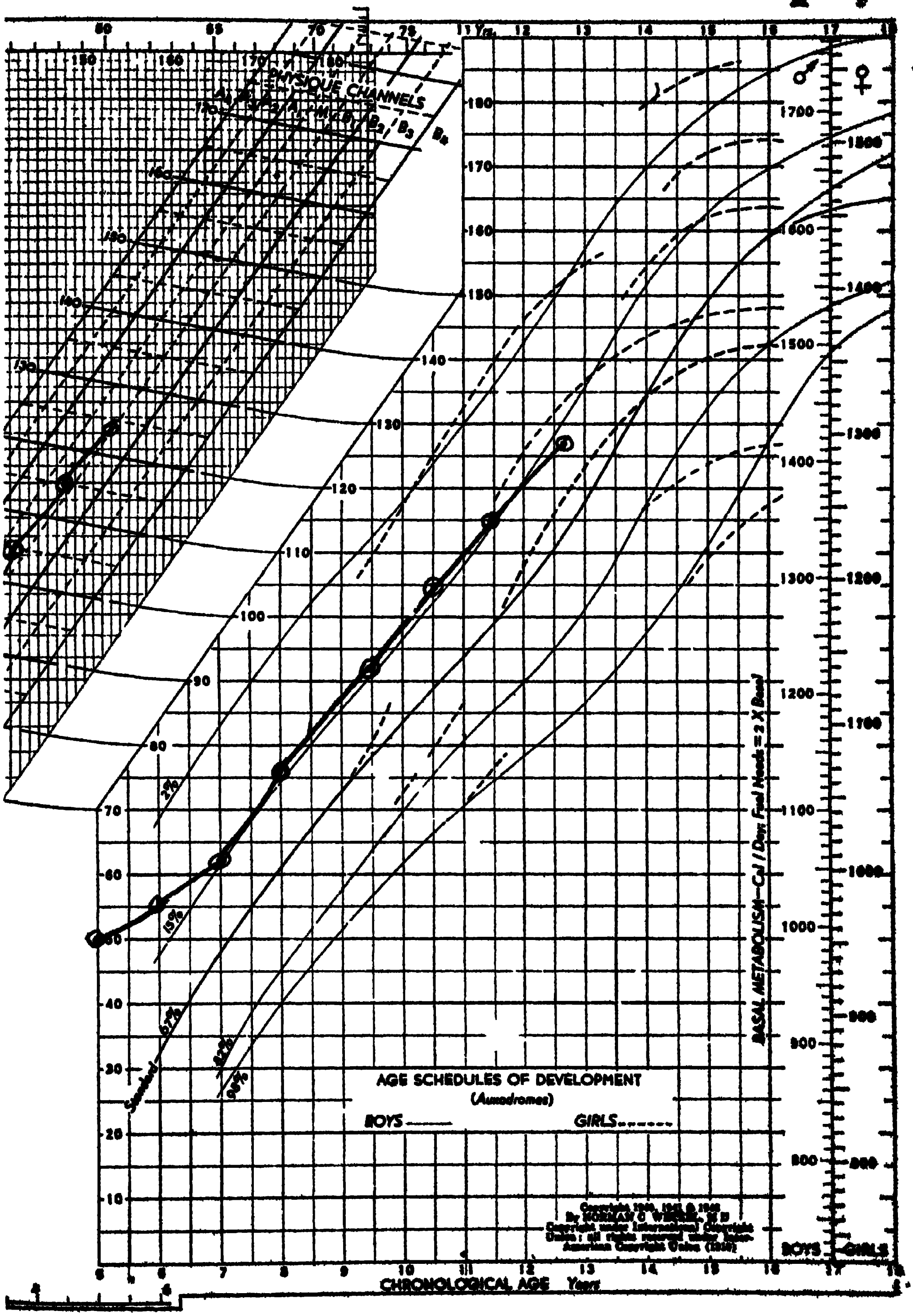
CLINICAL RATINGS

Gifted	A_4
Stocky	A_3 A_2
Good	A_1 M B_1
Fair	B_2
Borderline	B_3
Poor	B_4...

Fig. 4-6. The

ating PHYSICAL FITNESS
EVELOPMENTAL LEVEL and BASAL METABOLISM
Progress from Infancy to Maturity—

No.................



Wetzel Grid.

physique channels for each type. As long as a growing child advances steadily in his own channel, his height and weight are considered normal. But when a routine measurement of height and weight shows that he is veering out of his own channel, growth trouble is indicated This may be caused by prolonged emotional disturbance, glandular disturbance, serious malnutrition, or the early stages of some disease Wetzel's seven vertical physique channels represent different, normal body shapes from obese (A-3) through medium build (M) to very thin (B-3). The grids are printed on 9- by 12-inch file cards and can be easily and accurately kept current by the school nurse

Developmental levels are noted at the spot where weight and height are plotted For example, in Figure 4-6 the child 44 inches tall and weighing 48 pounds would fall at the developmental level 50 This level line number is then plotted on another graph (lower right corner) against age to determine whether or not the child is growing as fast as the norm for his age indicates In this case *level* refers to the measure of body size; each level represents a certain value of body surface By noting the level a child has achieved on the channel section of the grid and plotting it for a given age on the lower right-hand corner graph, a curve is obtained This is known as the child's *auxodrome* The horizontal scale is used for plotting ages, the vertical scale for plotting developmental level scores Auxodromes are lines running obliquely across the panel to show the percentage of children progressing at age schedules of development The auxodrome is blue for boys, red for girls, and black for the sexes advancing at the same rate The 67 per cent auxodrome tells whether the girl's or boy's progress is "normal," "advanced" or "retarded" The child is advanced when his curve lies above and to the left of the standard He is retarded when his curve lies below and to the right of the standard In this manner the pupil's rate of growth is determined

Although it may appear otherwise, the Wetzel Grid is not a complicated device to use Plotting scores and observing them over the years is not difficult for the average teacher Teachers and nurses are able to observe children regularly The thin child, for example, who suddenly begins to register his weight in the medium or stocky channel requires prompt investigation to determine the reason for the change In the case of a stocky child who drops to the thin territory on the grid, study is likewise necessary. Aside from presenting early evidence of deviation from normal growth, "the grid properly interpreted may save needless parental worry over their child, who, regardless of mistaken effort to increase his weight, remains thin, and over the stocky child who, in spite of needless exhortation to eat less, remains stocky."[85] Each child presents a graph that spells "normalcy" for him, and he is not forced to conform to the growth pattern of the group.

[85] "School Health Services," *Report of Joint Committee of the National Education Association and the American Medical Association,* Washington, National Education Association–American Medical Association, 1953, p. 37.

Still another means of appraising the growth and behavior characteristics of youth is the *Sheldon somatotype*.[36] The somatotype is a calculated estimate of body build. Persons who show heavy, fleshy structure (endomorphs) are quite different than those who are slim, small-boned, and delicate in structure (ectomorphs). Falling between these two extremes are the well-muscled, broad-shouldered, athletic persons (mesomorphs). For a discussion of somatotyping and the interesting relationship of health status and physical education skills to body build, see Chapter 13 Certainly, the extremes of body build yield considerable information about health characteristics and deviations[37]

Anthropometrical Screening Measurements

Since the days of Dr Dudley A Sargent, Dr. Edward Hitchcock, and Dr. William G. Anderson, health educators have been interested in the relationship between health status and various measures of body proportions, quantities of tissue, and muscular strength. Such items as vital capacity, pupil height-weight ratios, and the girth of various body segments have been examined by anthropologists and physical educators alike in order to devise a relatively easy-to-use measure for screening children for low fitness and unsatisfactory growth

A number of tables and manuals appeared between 1900 and 1925 determining the optimum weight and body proportions for a given height, sex, and age Montessori's *ponderal index*, of 1913, is still being used in some present-day research to give a quick picture of the varieties of body build.[38]

$$\text{Ponderal index} = \frac{\text{height}}{\sqrt[3]{\text{weight}}}$$

The height/cube root of weight ratio is reasonable when one considers that the body is assumed to be of constant density Thus the cube root of weight would yield a value proportional to the length of the side of a cube equal to the body in volume

In 1918, Pirquet, forced to do something about the nutritional requirements of the hungry Austrians, developed the *Pelidisi formula*.[39]

$$\text{Pelidisi formula} = \sqrt[3]{\frac{10 \times \text{weight (in grams)}}{\text{sitting height (in centimeters)}}} = 100\%$$

[36] William H. Sheldon, *The Varieties of Human Physique*, New York, Harpers & Brothers, 1940.

[37] Carl E Willgoose, "Health Characteristics of the Somatotypes," *The Physical Educator*, 12·3–5, October, 1955

[38] Maria Montessori, *Pedagogical Anthropology*, Philadelphia, Frederick A. Stokes Company, 1913

[39] William E. Carter, "The Pirquet System of Nutrition and Its Applicablity to American Conditions," *Journal of the American Medical Association*, 77:400, Nov. 12, 1921.

Here the properly nourished child rated about 95 to 100 per cent, the heavy, obese child up to 110 per cent, and the small, thin child as low as 88 to 94 per cent.

By 1920, the constitutional physiologist C. B. Davenport came out with his best index of build weight/height² Franzen, in 1929, with the adipose tissue calipers, studied the precision of a number of anthropometric measurements using modern statistical techniques.[40] Later Franzen and Palmer developed the *ACH index*, a battery of seven measurements: (1) hip width, (2) chest depth, (3) chest width, (4) height, (5) weight, (6) arm girth, and (7) subcutaneous tissue over the upper arm [41] The battery was later reduced to three tests: girth of upper arms (A), depth of chest (C), and width of hips (H).

After an intensive research project in which the factor analysis approach was used, McCloy recommended that schools use anthropometric measurements of height, weight, hip width, chest circumference, leg girth, width of elbow and knee, and girths of upper arm, forearm, and thigh [42] McCloy found that these elements were highly correlated with normal weight.[43] Few schools, however, have taken the time specifically to carry out the McCloy measurements More recent research in the area of body build, carried on by Sheldon and others, substantiates much of McCloy's work by noting the vast differences in appearance, health status, and general behavior present in the various types of human constitution

Among the many anthropometric devices for measuring growth and nutritional status are the Pryor Width-Weight Tables [44] Pryor was chiefly interested in nutritional status She took measurements of adolescent children every six months over a four-year period and demonstrated that during this period of rapid growth the bi-iliac diameter (pelvic crest width) is a useful measure in predicting body build Age-sex-width-weight tables were set up based on the following formula:

$$\text{Width-length index} = \frac{\text{bi-iliac diameter}}{\text{standing height}} \times 1{,}000$$

Norm tables are available from the Stanford University Press

Cureton devised an equation for predicting proper individual weight

[40] Raymond Franzen, *Physical Measures of Growth and Nutrition*, New York, American Child Health Association, Monograph no 11, 1929

[41] Raymond Franzen and George T. Palmer, *The ACH Index of Nutritional Status*, New York, American Child Health Association, 1934.

[42] Charles H McCloy *Appraising Physical Status The Selection of Measurements*, State University of Iowa Studies in Child Welfare, vol 12, Iowa City, State University of Iowa, 1936

[43] *Ibid.*, p. 32.

[44] Helen B. Pryor, *Width-Weight Tables for Boys and Girls 1 to 17 years; For Men and Women 13 to 40 Years*, Stanford, Calif., Stanford University Press, 1940; revised edition, 1947.

on the basis of the skeletal measurements of bone, muscle, and fat.[45] *Cureton's formula* has been used in the nutritional appraisal of junior high school boys. Here a skeletal index, a muscle girth index, and an adipose tissue index are combined. This was later improved by the substitution of the following formula:[46]

$$\text{Weight (lb)} = 9.09\ (\text{ankle girth, in.}) + 5.01\ (\text{minimum chest girth, in.}) + 4.12\ (\text{hip width, in.}) + 1.61\ (\text{height, in.}) - 249.55$$

Screening for Nutritional Disturbances

The relationship of malnutrition to poor posture, height, weight, general body build, and physical growth characteristics is so definite that many of the measures already referred to could easily be classified primarily as nutritional measurements. Although there are a number of weaknesses in using height-weight figures to indicate nutritional status, they have, nevertheless, been useful over the years as a crude screening instrument. Thus the early age-height-weight tables of Wood and Baldwin, distributed widely by the U.S. Department of Education, Public Health, and Labor, were useful in their day[47] The ACH index was also used in studying nutrition. Nearly every index that relates to body build, including the Meredith tables with the five zones for height and weight, the Wetzel Grid and its seven channels, Pryor's Width-Weight Tables, the Pelidisi formula, Cureton's formula, and Sheldon's somatotype, relates in some manner to individual pupil nutritional status

For the average teacher in the classroom, the health guidance counselor, and the physical instructor a simple check list of readily observable signs of malnutrition can be more helpful than the complicated screening devices. Certainly the check list or rating sheet supplements other screening procedures The signs and symptoms of good and poor nutrition that can be combined in one check list are outlined in some detail by Bogert.[48] In *Roberts' Nutrition Work with Children*, Martin sets forth a limited and concise check list for appraising the nutrition of school children.[49] This was developed by Dean Elizabeth N Todhunter, of the University of Alabama.

Screening for Mental and Emotional Adjustment

The relationship of mental health to successful adjustment in school is such that school personnel are frequently interested in appropriate screen-

[45] Thomas K Cureton, *Physical Fitness Appraisal and Guidance*, St. Louis, The C. V. Mosby Company, 1947, pp. 141–142

[46] *Ibid*, p 145.

[47] T. D Wood and B T Baldwin, *Height-Weight-Age Tables*, New York, American Child Health Association, 1918

[48] L. Jean Bogert, *Nutrition and Physical Fitness*, 6th ed., Philadelphia, W. B. Saunders Company, 1954

[49] Ethel Austin Martin, *Roberts' Nutrition Work with Children*, Chicago, University of Chicago Press, 1955, p. 363.

ing measures. No attempt will be made here to review the numerous tests of mental and emotional adjustment Because health and physical education teachers attempt to build individual social efficiency, Chapter 15 has been devoted to the topic A number of tests, inventories, and surveys of person-

RESEARCH LABORATORY OF HUMAN NUTRITION
UNIVERSITY OF ALABAMA

LIMITED CHECK LIST FOR NUTRITION
APPRAISAL OF SCHOOL CHILDREN

School __________ Observer __________ Date __________

Name of child __________ Age _____ Sex _____

	Desirable Characteristics		Undesirable Characteristics	
Head	up	wellshaped	forward	
Shoulders	even		rounded	
Back:	curves not exaggerated		much curved	
Chest	broad		shallow	
	deep		pigeonchested	
Limbs:	straight		bowed	
			knees touching	
Abdomen	flat		protruding	
Skin	clear		dry or scaly	
	glow of health		unnatural pallor	
Eyes	clear		squint at bright light	
			circles below	
Hair	lustrous		dry	scanty
Teeth	even	well spaced	cavities	missing
Mouth	free from sores		sores at corners	
Expression	alert	happy	listless	tired
Walk	alert	full of life	no pep	
Repose	relaxed		restless	overactive
			irritable	fatigued

Fig. 4-7. Check List for Appraising Nutrition SOURCE Research Laboratory of Nutrition, University of Alabama Used by permission.

ality, behavior, and social acceptance give some indication of emotional health problems They are more meaningful when supplemented by teacher observations and tests of intelligence and achievement

Tuberculin Testing

In 1937, Frost pointed out that the tubercle bacillus was losing ground and "that the eventual eradication of tuberculosis requires only that the

present balance against it be maintained."[50] Tuberculosis may eventually be eradicated, but it is still with us. At the present rate there is a good chance it will be considerably reduced by 1977.[51] Therefore, schools should continue their efforts to detect children who show sensitivity to uniform doses of tuberculin. The diagnosis is based on the assumption that all cases of tuberculosis react to tuberculin. It is desirable to test persons of all ages, but this presents certain difficulties; the uniform testing of school children with a uniform dose of tuberculin appears to provide a satisfactory means of comparing tuberculin sensitivity in various areas.[52] This method has been encouraged by the American School Health Association in its school accreditization program.

The two most common tuberculin tests are the Mantoux and Vollmer patch. In the case of the Mantoux test, a small amount of tuberculin is injected into the skin. The physician examines the spot after forty-eight to seventy-two hours to see if a hard blob at least 6 millimeters in diameter (pink or red) has been formed. If so, the subject is judged "positive." In the Vollmer patch test a tuberculin-impregnated patch is attached to the skin. In the "positive" person the reaction is similar to that in the Mantoux test. The Mantoux test, however, appears to be a better test because its response is prompt and more definite.

Fluoroscopy and X-ray photography are also used to diagnose tuberculosis. Very often this is used to follow up "positive" findings by discovering where the active infection lies.

The School Health Program as a Whole: Its Evaluation

Of major concern to the school administrator, as well as to the director of the health service, is the impact of the *total* school health effort. This frequently necessitates special surveys and research projects. The whole school is scrutinized: the lunchroom, the playground, the gymnasium, sanitation procedures, etc. Standardized rating scales are employed. A reorganization is sometimes called for.

School health programs are seldom evaluated as a whole. It is a more common practice to examine parts of the program, such as health services, sanitary practices, secondary school safety, and school lunch. In fact, Neilson and Irwin found, after completing an elaborate study involving cities of all sizes, that there are common weakness in many programs.[53] There was found

[50] W. H. Frost, "How Much Control of Tuberculosis?" *American Journal of Public Health,* 27:756-759, August, 1937.

[51] Floyd M. Feldmann, "How Much Control of Tuberculosis 1937–1957–1977?" *American Journal of Public Health,* 47.1235–1241, October, 1957.

[52] Michael L. Furcolow, "On the Usefulness of the Tuberculin Skin Test," *American Journal of Public Health,* 46:1071–1080, September, 1956.

[53] Elizabeth A. Neilson and Leslie W. Irwin, "Analytical Study of School Health Practices in the United States," *The Research Quarterly,* 29:417–458, December, 1958.

a lack of coordination between school health services, physical education, and health education. There were also weaknesses in school health councils, parent conferences, school health records, follow-up procedures, and qualifications of personnel. This indicates a need for more comprehensive appraisal of the effectiveness of the total program. One might ask, "How does the health program work in the school and community?" or "What is the status of the program?"

Evaluation Instruments. A number of evaluation instruments have been prepared to help school administrators and health directors appraise their health efforts Such instruments are useful in analyzing present programs as well as initiating new ones. A number of state education departments have developed excellent check lists and rating scales for use in this area. They are sometimes incorporated in the state course of study in health education. They may be used as part of a self-survey or by specialists brought into the community from a university or state education department. Some forms and inventories are standardized and cover a wide range of topics of health services, school environment, and health instruction In Illinois, for example, a three-part inventory is used to gather opinions not only from teachers and administrators but from parents and pupils too, with subsequent discussion of the findings.[54]

The following list includes the better-known devices for appraising the school health program Most of it was compiled through the joint efforts of the American Association for Health, Physical Education, and Recreation, the American Public Health Association, and the American School Health Association

American Council on Education, *Evaluative Criteria: Health and Safety (D7)*, The Co-operative Study of Secondary School Standards, Washington, 1950

American Public Health Association, "Suggested Standards for Health Services in Secondary Schools," *American Journal of Public Health,* 42:1102 1111, May, 1952.

American School Health Association, "Appraisal Form for Evaluating School Health Services," *Journal of School Health,* 18:1–12, January, 1948

California State Joint Committee on School Health, Subcommittee on School Environment, *Checklist for Healthful and Safe School Environment,* California State Department of Public Health, San Francisco, Calif , 1957

Cushman, Wesley, *School Health Instruction Survey Form,* The Ohio State University, Columbus, Ohio.

Dearborn, Terry H , *A Check List for the Survey of Health and Physical*

[54] Illinois Secondary School Curriculum Program, *Concensus Study: Number 4, Inventories A, B, and C,* Springfield, Ill , Illinois Department of Public Instruction, March, 1952.

Education Programs in Secondary Schools, Stanford, Calif., Stanford University Press, 1951.

Hicks, Dora A., *Evaluation Survey for Improvement in Health Education,* University of Florida, Gainesville, Fla., 1953.

Illinois Curriculum Program, *Concensus Study, Number 4: Inventory A. What Do You Think about Our Schools' Health Education Program? Inventory B. In What Respects Should We Strengthen Our Schools' Health Education Program? Inventory C. What Should We Do to Strengthen Our Schools' Health Education Program?* Springfield, Ill., Department of Public Instruction, March, 1952.

Knutson, Andie L., "Evaluating Health Education," *Public Health Reports,* 67:1–6, January, 1952

LaPorte, William A., *Health and Physical Education Score Card No. 1, for Elementary Schools, Health and Physical Education Score Card No. 2, for Junior and Senior High Schools and Four Year High Schools,* Parker and Company, 241 E. 4th St., Los Angeles 13, Calif.

Minnesota Department of Health and the State Department of Education Evaluative Schedule, *School Health Manual,* 2d ed., St Paul, Minn., State Department of Education, 1950

Mississippi State Board of Health and Department of Education, *Evaluation of Health Services on the School Health Program,* Jackson, Miss., 1952

Oregon State College, *A School Health Program Evaluation Scale,* Corvalis, Ore., 1955

Smith, Sara Louise, *Evaluation of School Health Program by Classroom Teachers,* rev form, The Florida State University, Tallahassee, Fla., 1959

Texas Education Agency, *A Checklist Appraising the School Health Program,* Bulletin 519, Austin, Tex., State Education Department, 1955.

University of the State of New York, State Education Department, *Evaluation of School Health Education,* Albany, N.Y., December, 1952.

Some of the better evaluation instruments designed to measure the effectiveness of a school health program simply call for a straight "yes" or "no" answer. This kind of instrument lends itself to quantitative treatment but may not tell too much about a particular detail. The Texas survey form, *A Checklist Appraising the School Health Program,* is an example of an instrument that can serve as a guide to a good program. It may be reproduced and is not at all difficult to complete. It is divided into eight parts, covering school health service, school nutrition program, healthful school environment, health of the school personnel, hygiene of the instruc-

tional program, health instruction, health coordination, and organization and administration of the school health program. The following example comes from the school health service section:

M. Dental health	*Yes*	*No*
1. A dental examination is given to the pre-school child	______	___
2. A yearly dental examination is given to all school children	______	___
3. Provision is made for the correction of dental defects through		
a. referral to school dental clinic	______	___
b. referral to dentist in private practice	______	___
4. Teachers of all grades promote dental health	______	___
5. Follow-up to ensure correction of dental defects is done by		
a. school nurse	______	___
b. teacher		___

Another example of a useful instrument is the *School Health Manual* of the Minnesota Department of Health and State Department of Education This schedule has been widely used for a number of years It may be mimeographed or otherwise duplicated for school use without permission It affords a rather complete appraisal of the total health program

A survey instrument or rating scale that throws more light on a given topic than the ordinary "yes" or "no" answer is illustrated by Sara Louise Smith in her form, *Evaluation of School Health Program by Classroom Teachers.* The following questions are from section 3, on Health Practices and Attitudes.

2. What is the attitude of children toward staying at home when there is danger of spreading disease by coming to school? ______

3. What is the attitude of parents toward keeping sick children at home? ______

Answers to such questions are sometimes difficult to tally for statistical purposes, but they are individually meaningful in surveying classroom attitudes.

Still another type of evaluation instrument is being used with considerable success in the Northwestern part of the country. This is *A School Health Program Evaluation Scale,* developed at Oregon State College, in 1955. This scale, appraising the various aspects of the total health program, permits numerical scoring of 1,000 possible points. Points are distributed in three areas as follows:

Part I	School Health Services	350 points	______
Part II	Health Instruction	400 points	______
Part III	Healthful School Living	250 points	______
		Composite score	______

The following questions from Part I serve as an example of how some of these points are set up in the scale:

A. Health Appraisal
   1. Frequency of health examinations
      a) Entering pupils examined—90–100% (15 pts.), 80–89% (12 pts.), 70–79% (9 pts.), 50–69% (6 pts), 20–49% (3 pts.) . . . . (15 pts.) ______
      b) New pupils entering the school system examined—90–100% (5 pts.), 80–89% (4 pts.), 70–79% (3 pts.), 50–69% (2 pts.), 20–49% (1 pt.) . . . . (5 pts.) ______
      c) Pupils examined at least once through grades three to ten—90–100% (5 pts), 80–89% (4 pts.), 70–79% (3 pts), 50–69% (2 pts), 20–49% (1 pt.) . (5 pts.) ______
      d) Pupils referred by teacher, or nurse, examined—90–100% (10 pts), 80 89% (8 pts.), 70–79% (6 pts.), 50–69% (4 pts), 20–49% (2 pts) . (10 pts.) ______
      e) Interscholastic athletic participants examined—90-100% (5 pts.), 80–89% (4 pts), 70–79% (3 pts.), 50–69% (2 pts.), 20–49% (1 pt.) . . (5 pts.) ______
      f) Pupils having injuries at school or serious illness necessitating absence of five days or more from school examined—90–100% (5 pts.), 80–89% (4 pts.), 70–79% (3 pts), 50–69% (2 pts), 20–49% (1 pt.) . (5 pts.) ______
      g) Pupils tested with audiometer every three years or more often (Elementary Schools only)—90–100% (5 pts), 80–89% (4 pts), 70–79% (3 pts), 50–69% (2 pts.), 20–49% (1 pt) . (5 pts.) ______

The Value of the Self-survey. Results of survey techniques provide a base line from which school health programs can be reoriented The survey will tend to be more valid and useful if the following points are considered:

1. Use qualified personnel to carry out the survey. These may include persons from the schools, local health department, and medical and dental professions; PTA members; parents; and others. This kind of endeavor leads to a mutual understanding of problems, which, in turn, leads to the formulation of recommendations for solving the problems Such a procedure is frequently more productive in getting action than the authoritative opinion of experts.

2. Do not repeat the survey too often, for the repeat survey lacks the enthusiasm and motivation of the original and tends to degenerate into a routine completition of questionnaires and forms.

3. In developing standards, remember that each school and school system has its own peculiar needs and problems. It is necessary, therefore, to develop check lists to meet the local conditions.

4. When comparing schools, use statistical procedures carefully. When applying statistical methods to school health program data, one should be

certain that the data have been collected accurately and interpreted properly. Statistics measure the volume and type of work done. They do not usually measure quality but may lead to a qualitative judgment. For example, one school may show that a large number of children have been referred for malnutrition, whereas in another school the ratio is only half as great. This may lead one to conclude that the school lunch or health-teaching program is more effective in one school than another. In reality, however, the difference in school referrals may be due to socioeconomic level of the pupils. In such a case the school surveys were not carried out entirely on a comparable basis.

5. Subject the results of the survey to wide discussion. Survey figures prove fruitful and enlightening when school and community persons begin to talk about them and ask further questions. This kind of action promotes constructive change and is an example of democratic action at its best.

SELECTED REFERENCES

Blum, Henrik L.: "Design and Evaluation of a Vision Screening Program for Elementary School Children," *American Journal of Public Health,* 49: 1670–1681, December, 1959.

Bogert, L. Jean: *Nutrition and Physical Fitness,* 6th ed., Philadelphia, W. B. Saunders Company, 1954, p 535.

Cureton, Thomas K.: *Physical Fitness Appraisal and Guidance,* St. Louis, The C. V. Mosby Company, 1947, chap. 5.

Foote, Franklin M., and Marion M Crane: "An Evaluation of Vision Screening," *Exceptional Children,* 20:153–161, January, 1954.

Knutson, A. L.: "Evaluating Progress in Health Education," *Journal of Health, Physical Education and Recreation,* 28:21–27, May–June, 1957.

Sixth National Conference on Physicians and Schools: "Youth Fitness," *Journal of Health, Physical Education and Recreation,* 29:22–26, February, 1958.

Wetzel, Norman T.: "The Simultaneous Screening and Assessment of School Children," *Journal of Health and Physical Education,* 13:200–205, December, 1942.

Wheatley, George M., and Grace T. Hallock: *Health Observation of School Children,* New York, McGraw-Hill Book Company, Inc., 1956.

Yankauer, A., and R. A. Lawrence: "A Study of Periodic School Medical Examinations," *American Journal of Public Health,* 45:71–75, January, 1955.

Yoho, Robert: "Fitness through Health Education," *Journal of Health, Physical Education and Recreation,* 28:18–22, September, 1957.

Young, Majorie A. C.: "The Brookline School Health Study," *Journal of School Health,* 31:47–54. February, 1961.

CHAPTER 5

Measures of Health Knowledge, Attitudes, and Habits

One of the most difficult items to evaluate is measurement of health knowledge, health attitudes, and health habits Man behaves the way he does because of many variables His actions and attitudes of the moment are related to his understandings ambitions, feelings, and interests Intelligent health behavior is frequently a quiet, subtle reaction that is hard to identify, it may also be an open, readily observable response And there can be a "sleeper effect" Pupils learn right from wrong what to eat, how to sleep, recreate, socialize, etc, but it may be several months or years before they see the personal value of putting this knowledge into practice Even favorable test results may be questioned, for some pupils are able to show nicely what they know or appreciate on paper, but their knowledge is frequently divorced from day-to-day living Measurement results in this area, therefore, must be carefully weighed to ascertain improvement correctly.

This chapter is devoted to the outcomes of health instruction, which can be thought of as elusive and uncertain Because of this, there is a distinct difference between knowledge tests, on the one hand, and tests of attitudes and habits, on the other It is true, however, that the two types of test complement each other For example, a student's habits and attitudes toward a certain health problem indicate to some extent his level of understanding As true as this may be, "it cannot be emphasized too much that the hazard of all health teaching is the tendency to follow traditional procedures and to measure progress by the ability to recite the topics assigned "[1]

HEALTH TESTS BY GRADE LEVEL

Elementary School

Brewer-Schrammel Health Knowledge and Attitude Test
California Tests in Social and Related Sciences
Crow-Ryan Health and Safety Education Test

[1] University of the State of New York, *Health Teaching Syllabus for the Junior and Senior High Schools,* Albany, N Y, The University of the State of New York Press, Bulletin no 1269, May, 1945, p 13.

Franzen-Derryberry-McCall Health Awareness Test
Gates-Strang Health Knowledge Test
Getchell Health Knowledge and Attitude Test
National Achievement Tests: Health Test (Speer-Smith)
Public School Achievement Test (1) Health
Survey Test in Health Education (Boyer)

Secondary School

Begbie Health Knowledge and Attitude Test
Brewer-Schrammel Health Knowledge and Attitudes Test
Byrd Health Attitude Scale
Gold's New Test in Health Knowledge
Johns and Juhnke Health Practice Inventory
Kilander Health Knowledge Test
Mayshark Health and Safety Attitude Scale
Myers Safety Attitude Scale for the Seventh Grade
National Safety Council Tests
Nehr Health Inventory
Shaw-Troyer Health Education Test
Stradtman and Cureton Physical Fitness Knowledge Test
Trusler-Arnett Health Knowledge Test
Veenker Health Knowledge Test for Seventh Grade

College

Bridges Health Knowledge Test
Byrd Health Attitude Scale
Cushman-Bennett Checklist of Health Problems
Dearborn Health Knowledge Test
Johns and Juhnke Health Practice Inventory
Kilander Health Knowledge Test for College Students
Leonard and Horton Inventory of Points of View Related to Health
Shaw-Troyer Health Education Tests
Trusler-Arnett Health Knowledge Test

Objective Tests of Health Knowledge and Understanding

A test that will measure not only the amount of health knowledge a pupil possesses but also the degree of his *understanding* of health problems is a particularly useful instrument. Moreover, the measurement of understanding goes beyond simply tallying health facts, because the pupil who understands is able to act, feel, and think intelligently, thus showing intelligent health behavior.

Too many teachers through the years have simply handed out facts to their pupils and have received the facts back at examination time. Thousands of school children can spell, write, and recite the names of the vitamins

and minerals found in specific foods; they can name the teeth and diagram how the blood circulates in the body. Yet many of them do not choose foods wisely; they don't own a toothbrush or appreciate the relationship of a sound heart to personal well-being in work and play. In short, they are rated as good students of health because they have given the teacher what she wanted to hear—health facts and figures.

Health knowledge and understanding is manifested in a number of ways. Intelligent health behavior in the school and community is illustrated by such concrete items as the increase in children having dental work completed, the number of improved habits of milk drinking, cleanliness, posture, personal appearance, etc., increased sale of nutritious foods at the cafeteria, reports on changed habits announced by parents, increased numbers of parents attending school health examinations of their children, and increased percentage of physical defects remedied and questions asked by pupils.

Another problem in measuring the results of health teaching is one of evaluation instruments Although local knowledge tests are not difficult to construct, good homemade tests are harder to make. This is further complicated because even if a pupil scores poorly in understanding and behavior, this does not mean that his behavior will not change as he matures and applies the knowledge gained previously in the classroom. Even when there are fairly good tests, their accuracy is reduced because some teachers fail to determine what a pupil or class knows about a topic at the beginning of the study In other words, some pretesting or early appraisal is necessary before a particular unit of work is started Shaw points out that this pretesting need not be elaborate, and it need not be a written examination: "even impromptu discussions at the beginning of a unit of study may partially serve this function"[2] The teacher's objective is simply to find out what the class knows and how it feels about the topic *before* instruction begins.

There follow comments on the better standardized health knowledge tests available for elementary school, secondary school, and college use

National Achievement Tests: Health Test.[3] This test, suitable for grades three to eight, was developed in 1938 and improved upon over the years by Speer and Smith. The test has a number of questions designed to test the students' ability to select the best health habits; ability to make comparative judgments; understanding of health causes and effects; and knowledge of health facts There are two current tests, Form A and Form B. Test reliability is given as .84. It takes about fourty minutes to complete. Norms are established by grades, and the publisher makes an attempt to keep the norms current by soliciting test scores. The four parts of the test together give a picture of the individual pupil's understanding of health. The four

[2] John H. Shaw, "Evaluation in the School Health Instruction Program," *American Journal of Public Health*, 47:582–585, May, 1957.

[3] Robert K. Speer and Samuel Smith, *National Achievement Tests: Health Test*, Rockville Center, N.Y., Acorn Publishing Company, Inc., 1949.

parts are Part I, Recognizing Best Habits; Part II, Health Comparisons; Part III, Causes and Effects; Part IV, Health Facts.

Health Education Test: Knowledge and Application.[4] To learn whether or not a pupil has grown in health knowledge and the ability to think critically in health matters, Shaw and Troyer developed a test that combines basic knowledge (Part I) with application (Part II). It is designed for grades seven to twelve and college. Both specific questions and norm tables have been carefully validated and kept current. Reliability is .92. In a testing time of forty-five minutes, secondary school and college students are presented with problem situations, the answers to which provide the instructor with a measure of the pupil's application of factual knowledge. Each form of the test consists of 100 multiple-choice and true-false questions. Here the student is asked to select the best answer:

18. Girls' participation in strenuous physical activity

F. Should be similar to boys'
G. Should be discouraged because it is unladylike
O. Is detrimental to health
L. Is desirable if carefully planned and supervised

Here is an example of a problem useful in determining application of knowledge:

PROBLEM A. John is 17 years old and a senior in high school. He lives in a northern state on a certified dairy farm. On this farm there is much energy consuming work between 5 A.M. and 7 P.M. John works one hour before school and two hours after school each day. For breakfast he usually eats bread, butter and jam, fried ham and eggs, fried potatoes or mush, pie, and coffee or raw milk. A typical evening meal consists of meat, potatoes (mashed or baked), bread, butter and jam, two canned vegetables, canned fruit and cake, coffee or milk. The noon meal for the family is similar to the evening meal, but during the winter months John eats lunch at the school cafeteria where he usually chooses· meat, potatoes, bread and butter, pie, and chocolate milk.

John usually gains 15 to 25 pounds during the winter, but loses most of this weight during the summer when he is working full-time. Other members of John's family are not more than 10 per cent overweight Bill, John's brother, says he keeps his weight down by smoking a package of cigarettes each day.

T F 61. John probably gets too much to eat for breakfast.
T F 62 John probably gets more calories than he needs for lunch at school.
T F 63. It would be well for John to eat some fresh fruit when he returns from school.
T F 64 John is safe in drinking raw milk.
T F 65. John's diet should include cod liver oil during the winter months.
T F 66. Pie is not an appropriate food for breakfast.
T F 67. The diet of John's family would not be appropriate for a family of business or professional people.

[4] John H. Shaw and Maurice E. Troyer, *Health Education Test: Knowledge and Application*, Rockville Center, N.Y., Acorn Publishing Company, Inc., 1956.

T F 68. The eating of potatoes three times a day provides too much carbohydrates in the diet of this family.

T F 69. Bill has a desirable method of controlling his weight.

Gates-Strang Health Knowledge Test.[5] The latest revised edition by Arthur I. Gates and Ruth M. Strang contains sixty multiple-choice items arranged for the elementary grades. There are appropriate grade level norms and three forms of the test: Forms A, B, and C. Test reliability ranges from .74 to .86.

Franzen-Derryberry-McCall Health Awareness Test.[6] This instrument was validated against health status and was designed to measure a combination of knowledge and attitudes. A battery of multiple-choice and true-false items is given to elementary pupils in grades seven to nine. The assumption in this test is that the healthy person who is at least free from easily observed major defects and diseases is one who possesses the knowledge and basic understanding of how to keep well. Although this assumption can be readily challenged (for many variables have a bearing on health status in addition to knowledge), the test is a useful one when questions are kept up to date. Test reliability ranges from .93 to .95.

Crow-Ryan Health and Safety Education Test.[7] The test is designed for grades four to six. The purpose of Part I is to measure the pupil's knowledge of good health and safety habits. A series of multiple-choice questions about common health and safety activities are set up for quick recognition. Part II tests the pupil's knowledge on cause and effect in relation to health and safety Part III is concerned specifically with facts about health and safety. For instance, one question is about the particular vitamin necessary for healthy eyes. The pupil is asked to choose between (1) vitamin B, (2) vitamin A, and (3) vitamin I . Part IV is concerned with the actual application of health and safety rules; there is a short problem, which the student must understand in order to make a correct choice from three possible answers. Here is an example from this intermediate grade level test [8]

PROBLEM 9. Mary does not eat breakfast but buys candy on her way to school which she eats during classes. Eleanor eats a big meal in the morning and a big meal in the evening but does not eat lunch. Julie has orange juice, cereal and milk for breakfast, eats a light lunch and supper consisting of meat, vegetables, a light dessert and milk. The best eating habits are practiced by

A. Mary B. Eleanor C. Julie

[5] Arthur I. Gates and Ruth M. Strang, *Gates-Strang Health Knowledge Test,* rev. ed., New York, Bureau of Publications, Columbia University, 1945.

[6] Raymond H. Franzen, Mayhew Derryberry, and William McCall, *Health Awareness Test,* New York, Bureau of Publications, Columbia University, 1946

[7] Lester D. Crow and Loretta C Ryan, *Acorn National Achievement Test, Health and Safety Education, Grades 3 to 6,* Rockville Center, N.Y., Acorn Publishing Company, Inc., 1956.

[8] *Ibid,* p. 8. Used by permission.

In the Crow-Ryan test, as in the Shaw-Troyer test, the publisher makes an effort to keep the norms up to date. Norms are established for each part of the test as well as for the total score for each grade.

Kilander Health Knowledge Test for High School Students.[9] Achievement tests were developed for senior high school pupils so the outcome of instruction in health could be evaluated Although it is strictly a health knowledge and understanding test, it is nevertheless useful in noting some personal attitudes and practices related to health. Kilander constructed the test in 1936[10] Questions were revised in 1950 and 1954. Few written health tests have been given so extensively to American school children. Its success seems due to ease of administration and simple scoring key. In addition, most of the seventy-five questions in the two forms, A and B, are the kind that do not easily change in meaning from year to year The test validity was established according to rigid standards The test items were constructed only after a thorough analysis of varied instructional materials and authoritative pronouncements in the field of health and health education Most elements measured can be justified by frequency of inclusion in commonly used textbooks and by expert judgment as to importance Reliability is 80 to 83 To make the test more useful, an expectancy table has been prepared, indicating the level of achievement on the Kilander Health Knowledge Test associated with varying levels of IQ for a random sample of students in the normative group Thus individual pupil achievement can be related to learning aptitude. Two sample questions follow:

67 Which one of the following chemical salts, when found in drinking water or applied directly to the teeth, seems to help reduce tooth decay?

1 Chlorides 2 Fluorides 3 Sulphates 4 Nitrates

75. Most people who are overweight are so because they

1 Exercise too little
2. Have inherited a tendency to be overweight
3. Have an underactive thyroid gland
4. Eat too much

Kilander Health Knowledge Test for College Students.[11] This test is designed to measure the extent of a college student's knowledge and understanding of matters pertaining to health. It is an eight-page test of 100 multiple-choice questions about personal health, nutrition, community

[9] H. F. Kilander, *Kilander Health Knowledge Test, Forms A and B,* Yonkers, N Y, World Book Company, 1952.
[10] H. F Kilander, "Health Knowledge of High School and College Students," *The Research Quarterly,* 8 32 36, October, 1937
[11] Frederick H. Kilander, *Kilander Health Knowledge Test for College Students,* Frederick H. Kilander, 33 Colonial Terrace, East Orange, N.J., 1936; 4th ed., 1958.

health, sanitation, communicable diseases, safety, first aid, family living, and mental health. Included are misconceptions, superstitions, and common errors. Validity was determined empirically from state courses of study and textbooks in health. Reliability was .83. The test can be used with high school students. Norms are available and are based on over one hundred thousand individual scores. Here are some sample questions.

1. The World Health Organization, known as WHO, is:

1. An agency of the old League of Nations
2. An independent international agency working closely with the UN
3. An agency of the United Nations
4. A loose international federation which includes most countries but not Russia and its satellites

13. Doctors generally consider bottle feeding of infants as compared to breast feeding to be:

1. Better 2. Just as good 3. Poorer 4. Debatable

26. The best method today of reducing the death rate from cancer is by:

1. Early diagnosis
2 Repeated use of radium and x-ray
3. Improvement in one's general health

79 According to present scientific knowledge, which one is entirely attributed to heredity.

1. Cancer 2. Excessive weight 3. Color-blindness 4. Anemia

Dearborn Health Knowledge Test.[12] This test provides a valid, reliable, and comprehensive measure of achievement and diagnosis of personal health knowledge. It is designed chiefly for college students but can be used with informed adults and high school seniors who have been exposed to a full semester course in personal hygiene. There are 100 multiple-choice questions set up in eleven areas:

Social and biological background	Reproduction and heredity
Nutrition and diet	Prevention and control of disease
Excretion and cleanliness	Hygiene of eyes, ear, and teeth
Exercises and body mechanics	Hygiene of environment and the use of medical care
Fatigue and rest	
Mental hygiene	

The test was developed from 1,200 test items in personal hygiene applied to 10,000 students by a number of instructors in colleges, universities, and

[12] Terry H. Dearborn, *College Health Knowledge Test,* Stanford. Calif., Stanford University Press, 1950; second printing, 1956.

junior colleges. Standardization was carried out on 2,000 students, with particular attention being given to the recommendations of health educators, an analysis of authoritative statements, and the preparation and application of four balanced trial forms. The mean coefficient of reliability for the test is .89. As to test validity, there is a moderately strong correlation between scores made on the test and the final grades given by the instructors at the close of the hygiene courses. Correlation coefficients in this instance varied between .64 and .87 for institutions making up 28 per cent of the sample. A norm table of percentile scores is available for junior college and college-university scores. As in the case of some other health tests, the users are asked to forward test scores to the publisher so that the norms can be extended. This is a good practice, providing the directions for giving the test are carefully adhered to and the resulting test scores appear to represent a true picture of the local testing experience. Here are two sample questions about reproduction and heredity.

66. Menopause is a condition which normally occurs during

1. Puberty
2. Early adult life
3. Pregnancy
4. Middle age
5. Old age

67. Ovulation in women normally

1. Occurs about half way between menstrual periods
2. Follows fertilization of the ovum
3. Occurs immediately preceding menstruation
4. Occurs in the oviducts
5. Occurs in the uterus

Trusler-Arnett Health Knowledge Test.[13] Although this test was developed several years ago, the questions are of such a nature that they are not outdated. Considerable attention is devoted to standard health facts that should be known by high school and college students. Test reliability on Forms A and B is .86. Raw test scores are converted into percentile scores and compared against a percentile table. These scores can then be translated into school marks or letter grades.

Bridges Health Knowledge Test.[14] Bridges worked out a health knowledge test for college freshmen at Indiana University, in 1952, which consists of 100 best-answer questions. After reviewing an extensive list of health

[13] V. T. Trusler, C. E. Arnett, Jr., and H. E. Schrammel, *Trusler-Arnett Health Knowledge Test*, Emporia, Kans., Bureau of Educational Measurements, Kansas State Teachers College, 1941.

[14] Frank A. Bridges, *Health Knowledge Test for College Freshmen*, Rockville Center, N.Y., Acorn Publishing Company, Inc., 1956.

knowledge tests, Bridges developed one that emphasizes by percentages certain topics above others.

Nutrition	16 per cent
Emotional health	12 per cent
Exercise and rest	6 per cent
Narcotics and stimulants	5 per cent
Body functions	10 per cent
Social health	8 per cent
Personal health	13 per cent
Family living	6 per cent
Sense organs	4 per cent
Occupational health	3 per cent
Home nursing	2 per cent
Current health	1 per cent
Community health	11 per cent

The test was validated by a jury of experts and further checked by item discrimination and item analysis Bridges obtained a coefficient of reliability of 83.

California Tests in Social and Related Sciences.[15] There is a useful health and safety test in this series designed especially for elementary grades four to eight Section A is about simple principles of nutrition Section B is about health problems involving disease control, teeth, alcohol, tobacco, and similar items Section C is about safe living practices The percentile norms were built on a sample of close to twenty-five thousand cases The questions have satisfactory validity and reliability and are of the true-false and best-answer type

Getchell Health Knowledge and Attitude Test.[16] Although Getchell attempts to measure both health knowledge and health attitudes, the test is primarily one of health knowledge It designed for grades four to eight and is arranged in three parts true-false, multiple-choice, and matching. Reliability and validity is undetermined and norms are not available, but the questions are clearly written and can be quickly administered with a single test sheet

Veenker Health Knowledge Test for Seventh Grade.[16a] This instrument was developed for the purpose of measuring the extent of health knowledge possessed by seventh grade pupils Appropriate vocabulary burden, level of difficulty, and curriculum validity were primary factors considered in item formulation The test is easy to administer and to interpret

[15] Georgia Sachs and John A Sexton, *California Tests in Social and Related Sciences, Part III, Elementary Form AA,* Los Angeles, Calif , California Test Bureau, 1953.

[16] Vera Getchell, *Health Knowledge and Attitude Test,* Emporia, Kans , Bureau of Educational Measurements, Kansas State Teachers College, 1947

[16a] For a copy of Form FA, see article by C Harold Veenker, "A Health Knowledge Test for Seventh Grade," *The Research Quarterly,* 30:338–348, October, 1959.

Additional Tests of Health Knowledge. There are other usable health knowledge tests, but space will not permit an extended coverage of their strong and weak characteristics.

Kilander Tuberculosis Information Test:[17] This test consists of twenty multiple-choice questions related to tuberculosis. It can be used with high school and college students to determine what students know about this infectious disease. Continued testing has indicated an improvement in the amount of knowledge in this area.[18]

Sex Knowledge Inventory:[19] This is a battery of test items designed to measure knowledge about the function of the sexes. This is part of family life education

Public School Achievement Test (j) Health:[20] This is a test for grades four to eight arranged in a series of eighty-four test items. It is easy to give and easy to score. Although it was an early test, it is still useful. Nationwide norms are available.

Gold's New Tests in Health Knowledge:[21] This test was constructed for grades seven, eight, and nine and originally validated on a test population of 4,322 junior high school students. Reliability is .91, and grade norms have been established. The test measures knowledge about diet and dental health.

Survey Test in Health Education:[22] This test was designed for public school children in grades four, five, six, seven, nine, and ten. There are six forms in use.

Stradtman and Cureton Physical Fitness Knowledge Test:[23] This test was prepared to estimate the knowledge of secondary school boys and girls about the nature of physical fitness and its elements.

Manchester Unit Elementary Tests:[24] These are individual health tests covering grades five to eight.

Health Tests for Elementary School Children:[25] These are informally constructed batteries of tests that are useful in suggesting test items to the instructor. They are not standardized and have several weaknesses; yet they are of interest to look over before building a local test.

[17] Frederick H. Kilander, *Tuberculosis Information Test*, Frederick H. Kilander, 33 Colonial Terrace, East Orange, N.J., 1947, 2d ed., 1951.

[18] Frederick H. Kilander, "Knowledge about Tuberculosis Held by High School and College Students," *Journal of School Health*, 24:160-163, March 1954.

[19] *Sex Knowledge Inventory*, Durham, N.C., Family Life Publications, Inc., 1950.

[20] Jacob S. Orleans and Glenn A. Sealy, *Public School Achievement (j) Health Form 1*, Bloomington, Ill., Public School Publishing Company, 1928.

[21] Leah Gold, *New Tests in Health Knowledge*, Northfield, Minn., Carleton College Press, 1944.

[22] P. A. Boyer, *Survey Test in Health Education*, Philadelphia, Pa., Board of Education, 1940.

[23] Alan D. Stradtman and Thomas K. Cureton, "A Physical Fitness Knowledge Test for Secondary School Boys and Girls," *The Research Quarterly*, 20:52–56, March, 1950.

[24] Bureau of Tests and Measurements, Manchester College, North Manchester, Ill.

[25] Examples are found in *Grade Teacher*, 64:74–76, *February*, 1947; 64:67, December, 1947; 65:100–101, March, 1948; and 65:62–63, May, 1948.

Objective Tests of Health Attitudes

As already indicated, appraising pupil attitudes and habits is difficult. An attitude, by definition, is a problem in measurement to the average research worker in education. To be noticed at all, an attitude has to be expressed in verbal and/or observable behavior It is, as Bain has suggested, a relatively stable "overt behavior of an individual which affects his status."[26] The chief difficulty is that an expressed attitude or even an opinion given on a test may not be a true statement of the person's views. There is considerable evidence that there are individual differences in the degree of stability and constancy of opinions.[27]

The question may be asked, "How do we know for certain that a pupil possesses a particular attitude or not?" Probably the most direct method is to observe his behavior, if one can be certain of getting a reasonable sample of behavior and not simply what he desires the instructor to see Consistency helps A consistent point of view, opinion, or attitude is often fair evidence that true understanding is present In this connection two or more test questions worded differently and with a concealed approach may receive consistent answers

Attitude measurement is still a new field despite several decades of testing efforts and research findings Certainly health education, if it is to be more than a process of presenting facts to students, must conscientiously attack the problem of appraising health attitudes This, of course, must go beyond thoughts and opinions to the personal practices and daily habits of persons in general Mayshark stresses the need for more of this kind of measurement [28] Edwards expresses the same view when he says that "it is high time we stop talking theoretically about the value of health attitudes and try to be practical and measure them objectively "[29]

Making a study over a thirty-year period, Mayshark appraised a large number of so-called attitude tests and found that few were standardized and only three specifically measured attitudes [30] The three were (1) the work of Boyd on attitudes toward food practices, (2) the health attitude scale of Byrd, and (3) the Mayshark scale for measuring health attitudes of seventh graders Mayshark discovered that in a number of test instruments numerous questions called for a response primarily based on knowledge rather than attitude Moreover, several tests had not been refined, were without norms, and in some instances were asking questions that called for

[26] Read Bain, "An Attitude on Attitude Research," *American Journal of Sociology,* 33 940 943, May, 1928

[27] A S Barr, William H Burton, and Leo J Brueckner, *Supervision,* New York, Appleton-Century-Crofts, Inc., 1947, p 240

[28] Cyrus Mayshark, "Critical Analysis of Attitude Measurement in Health Education 1927-57," *The Research Quarterly,* 29 309–312, October, 1958.

[29] Ralph Edwards, "An Aproach to Health Attitude Measurement," *Journal of School Health,* 26:215 217, September, 1956.

[30] Mayshark, *op. cit.,* p. 310.

different answers according to what part of the country the pupil happened to be from.

It seems safe to say that teachers at all grade levels must give some attention to the measurement of health behavior as demonstrated through personal attitudes and practices. Furthermore, a study of available tests helps prepare for a more valid and useful test as time goes on. There follows a brief review of the more common attitude tests.

Brewer-Schrammel Health Knowledge and Attitude Test.[31] This is a 100-item true-false and multiple-choice test for elementary school and junior high school students. It purports to measure both knowledge and attitudes. If the test score is high, the assumption is that this reflects sound health attitudes. In the validation procedure the high-scoring junior high school students (out of a total of 287) were considered to have sound health attitudes because of a previous estimation of their personal health habits The major part of the test is a measure of health knowledge. There is considerable doubt, therefore, that this test is of much value for measuring attitudes Mayshark feels this way. Schneider supports his view.[32] Patty goes further and questions the validity and reliability of an attitude test that claims to measure pupils in several grades, from elementary through junior high school.[33] Test reliability ranges from .66 to .83.

Byrd Health Attitude Scale.[34] This test has probably been the most popular over the years The high school or college student is asked to answer 100 statements dealing with a health activity or policy by underlining one of five options. It can be used nicely for pretesting and postinstruction testing. Its chief purpose is to identify and evaluate individual pupil attitudes. It can be given in thirty minutes. Answers to questions are scaled from "strongly agree" to "strongly disagree" Points range from 100 to 500. There are no specific norms. Reliability is reported to be .95. Byrd makes a particularly significant point when he says[35]:

> There is some preliminary evidence that an individual must strongly agree with a favorable health practice if his behavior is to be modified . Individual scores should range above 400 points before we can expect attitudes to be transmitted into desirable health behaviors In other words, low scores are apt to be more significant than moderately high scores, because such ratings would indicate that there is not sufficient intensity of favorable attitudes to insure sound health behavior.

[31] John W. Brewer and H E Schrammel, *Brewer-Schrammel Health Knowledge and Attitude Test,* Emporia, Kans , Kansas State Teachers College, Bureau of Educational Measurements, Form A, 1935, Form B, 1939.

[32] Robert E Schneider, *Methods and Materials of Health Education,* Philadelphia, W. B. Saunders Company, 1958, p 366

[33] Willard W. Patty, "Blanket Health Tests Unsatisfactory," *Journal of School Health,* 20:258–260, November, 1950.

[34] Oliver E. Byrd, *Byrd Health Attitude Scale,* Stanford, Calif., Stanford University Press, 1940.

[35] *Ibid.,* p. 4.

The following examples are from the Byrd scale. Note the directions given.

READ EACH ITEM CAREFULLY AND *AT ONCE* UNDERLINE YOUR REACTION TO THE STATEMENT. Please work rapidly. Indicate the way you honestly *feel* about each statement. Be sure to answer every item.

3. The health department should leave all health matters to the local doctors.

Strongly agree[1] Agree[2] Undecided[3] Disagree[4] Strongly disagree[5]

20. A person should have a regular hour for going to bed.

Strongly agree[1] Agree[2] Undecided[3] Disagree[4] Strongly disagree[5]

40. The schools should increase their medical services to children.

Strongly agree[1] Agree[2] Undecided[3] Disagree[4] Strongly disagree[5]

55. Compulsory physical education in schools should be abolished.

Strongly agree[1] Agree[2] Undecided[3] Disagree[4] Strongly disagree[5]

Boyd Scale of Attitudes toward Desirable Food Practices.[36] Working at the University of Kentucky, Boyd developed two tests for measuring attitudes toward desirable food practices:

Test one An attitude questionnaire for scaling attitudes toward gardens, food storage, and well-selected diets A total of 68 opinion questions validated by the technique of equal-appearing intervals[37]

Test two: A free-association test where the pupil is asked to respond to a health "stimulus" word Here the individual's response is not affected by his knowledge of the right answer Also attitudes may be tapped while they are still below the threshold of awareness.

Boyd obtained a correlation of agreement between the attitude scale and the free association test of .31. He explained that the attitude questionnaire is designed to appraise such attitudes as can be represented by opinions which pupils endorse, whereas the free-association instrument probably measures certain attitudes during early growth stages when they are not yet matured to the point of awareness. Thus the two sets of scores might not be expected to have a high degree of correlation The free-association technique for measuring health attitudes suggests more research into its use in health education.

Begbie Health Knowledge and Attitude Test.[38] This test is practically a straight health knowledge test of the true-false, multiple-choice kind. Only a small percentage of the questions appear to measure beliefs or feelings.

[36] G. R. Boyd, *The Construction of an Instrument for Measuring Attitudes toward Desirable Food Practices*, Lexington, Ky., University of Kentucky, Bureau of School Service, 16:1–89, September, 1943.

[37] *Ibid.*, p. 18.

[38] Z. Begbie, *Health Knowledge and Attitude*, Emporia, Kans., Kansas State Teachers' College, Bureau of Educational Measurements, 1949.

Mayshark Health and Safety Attitude Scale.[39] Mayshark makes an excellent case for this test on behaviorism. It is an attempt to eliminate the unrelated and frequently ambiguous short questions that thwart the pupil from giving a valid response. The technique employed is a situation-response item. Reliability of the alternate forms is .93. It is economical of time, and there was found to be a statistically significant difference between the health attitudes of seventh-grade boys and girls. One of the strong points about this instrument is that it is aimed specifically at the seventh grade. It does not attempt to cover a wide range of grade levels and thereby sacrifice validity and reliability. Here are two examples.

Form A

5. Several younger children have been hurt on and around the playground swings. You would:

1. Ask them to leave while you are playing there.
2. Swing yourself and ask them at the same time
3. Let them look out for themselves.
4. Not swing, but watch the younger children.
5. Leave the swings when the younger children are there.

Form B

5. Your father drives you to school on his way to work If you were both late one morning, you would:

1. Tell him to be careful and forget the time.
2. Let him drive as he wished
3. Ask him to drive faster all the way.
4. Ask him to drive faster when it seemed safe to do so.
5. Suggest he slow down in traffic.

In this test health and safety attitude testing go hand in hand. Other measures involving safety attitudes are also considered in this chapter (see section on Myers Safety Attitude Scale for the Seventh Grade).

From a study of such tests, it is obvious that valid measurement of health attitudes is difficult. These tests, however, are quite useful, particularly if the teacher uses them in the instruction program to bridge the gap between knowledge and behavior. Students at all ages lack knowledge and have strange attitudes toward health. Pretesting often illustrates this point. Dearborn, for example, gave his college health test to 12,000 college, university, and junior college students and found significant results showing that health ignorance, both recognized and unrecognized by the student, was widespread.[40]

[39] Cyrus Mayshark, "A Health and Safety Attitude Scale for the Seventh Grade," *The Research Quarterly*, 27:52–56, March, 1956.

[40] Terry H. Dearborn, "Personal Health Knowledge of College Students before Instruction," *The Research Quarterly*, 29:154–158, May, 1958.

Health Inventories

The health inventory is very much related to attitudes. It seeks answers from students that require some degree of self-analysis. The object is not a high test score but the discovery of vulnerable health spots in the daily living of boys and girls The completed inventory, therefore, becomes a practical tool for the teacher to refer to when discussing health problems with individual pupils.

Leonard and Horton Inventory of Points of View Related to Health.[41] Here is a college-level inventory used to examine student actions based on certain fundamental viewpoints or attitudes influencing health behavior. It is also designed to examine the particular health problems or conditions about which students feel a need for secrecy and to study student beliefs about the reliability of sources other than a physician in solving health problems. Although not specifically intended for senior high school pupils, it appears to have value at this level, particularly in health counseling. It can also be used as a pretest for appraising health attitudes before instruction and for retesting after instruction Sutton, for instance, found that the inventory was quite useful in noting changes in opinions.[42] In one case he found that 304 students were asked to respond whether they felt a need for secrecy about a disease such as tuberculosis in their family There were 20.4 per cent for secrecy on pretesting This figure dropped significantly to 05.3 per cent on the retest

Johns and Juhnke Health Practice Inventory.[43] This is a revised form of the 1943 Johns inventory It is a reliable instrument for appraising the health practices of senior high school and college men and women as well as adult groups Norms are available covering grades eleven and twelve and for college freshmen and sophomores Topics include personal health; nutrition; dental health; physical activity and recreation; rest, sleep, and relaxation; prevention and control of communicable disease, prevention and control of chronic diseases, stimulants and depressants, mental health; family health; consumer health; community health; and safety education. Test questions are clearly worded and require the student to make one of five responses, which are numerically rated 1 to 5 The test is scored by allowing 1 point for questions answered "never"; 2 points for "rarely"; 3 points for "sometimes"; 4 points for "usually"; and 5 points for "always." The test can be used to pretest, to appraise the teaching effort, to gather pertinent informa-

[41] Margaret L. Leonard and Clark W Horton, *An Inventory of Points of View Related to Health—An Inventory for College Students,* Sacramento, Calif., State Department of Education, November, 1949

[42] C. Sutton, "An Appraisal of Health Attitudes and Practices of College Students," *The Journal of School Health,* 21 · 125–127, April, 1956.

[43] Edward B Johns and Warren L Juhnke, *Health Practice Inventory,* Stanford Calif., Stanford University Press, 1952.

tion for constructing a course curriculum, and to determine health practices of persons or groups.

This test is mechanically well constructed and comes in an eight-page booklet complete with scoring key forms. Here is an example.

Do you

10. Eat one or more servings of leafy, green, or yellow vegetables daily?

1. Never 2. Rarely 3. Sometimes 4. Usually 5. Always

Neher Health Inventory.[44] A high school inventory designed to measure "What You Do About Health" (Part I) and "What You Know About Health" (Part II). Norms were built for grades nine, ten, eleven, and twelve on students sampled from many different racial, occupational, social, and economic groups. They are on both a percentile and a rating basis. Student scores are interpreted as "very low," "low," "average," "high," and "very high." Test reliability is .86. Although the original research for this inventory was conducted prior to 1942, there is little evidence to indicate that the test is not currently valid. The questions are well constructed and are not easily outdated. A Data Summary Form is provided for teacher use.

In this test health status is ascertained by such questions as "Do you wake up in the morning feeling tired?" "Do you find it necessary to use laxatives?" Here the student answers "frequently," "occasionally," or "never." There is also a series of questions designed to measure health practice Here a student is asked a question such as, "Do you go to bed to cure a cold?" He must answer either "yes," "sometimes," or "no" In Part II a statement based on information about health is given. The student must select one of five answers to illustrate what he knows about the statement.

Health Habits

A habit is much more than a show of knowledge or demonstration of an attitude. It is behavior resulting from a *balance* of knowledge, understanding, interest, and attitudes. When it can be observed, it can be measured more accurately. Unfortunately, the teacher cannot see the student all the time. The limited view of pupil behavior must be supplemented by interviews, well-chosen questions, rating scales and check lists, etc. If, for example, the pupil says that he sleeps well regularly, or has a daily bowel movement, or brushes his teeth after meals, the teacher must assume that he is telling the truth. If the teacher is sure that John Jones now sleeps well because he followed the teacher's advice and consulted his family physician, there is rather concrete evidence that health behavior (habits) has been favorably changed through the process of formal education.

Because of the nature of pupil interests, opinions, attitudes, and overt

[44] Gerwin Neher, *Health Inventory for High School Students,* Los Angeles, Calif., California Test Bureau, 1942.

behavior in general, it is more difficult to construct local tests of health habits than to develop tests of knowledge and application. Several measures are available, however, that can be used to note *tendencies* toward health habits and attitudes. It is usually true in any field that no one test can be trusted completely. Tests and measures, at best, are limited. Test results coupled with teacher observations and other factors improve the evaluation.

Clarke Health Habit Questionnaire.[45] This questionnaire was originally developed to be used with college students low in physical fitness. The questions are of such a nature that it can be used with both sexes at either the senior high school or college level. It is a most informative instrument, particularly if the student has ample time to complete it. Like many devices of this nature it is more valid when completed by the instructor during a regular pupil interview. Because the questionnaire is so useful, it is reproduced in full.

Health Interests

Sometimes a survey of the health interests of a group can be made. The teacher will find that making an inventory of health interests and using them to survey the class is most enlightening. With this information, the total instructional program can be made more meaningful. Classroom discussions, lectures, and conferences can be directed to the health questions and problems that are found. Loop and Tipton were quite enthusiastic about the survey they conducted because it not only gave them evidences of health practices and attitudes toward health but also provided them with the means for regrouping the topics in the hygiene course around real life situations.[46]

Persons tend to behave favorably toward things in which they are interested. Lantagne's study of the health interests of 10,000 secondary school children is a valuable contribution and one that shows that health interests of high school students can be measured with a high degree of reliability.[47]

Further information on the preparation and use of interest surveys is found in Humphrey's work on college men.[48] Another especially useful interest study was done with the public school children of Denver, Colorado, as a phase of the curriculum development program.[49] A health interest check list was used with pupils in grades four through twelve. Pupils were asked

[45] H. Harrison Clarke, *Health Habit Questionnaire (Revised)*, Syracuse, N.Y., Syracuse University Press, 1941; copies may be obtained from the author at University of Oregon, Eugene, Ore.

[46] Anne S. Loop and Anne B. Tipton, "A Health Survey of Hunter College Freshmen," *The Research Quarterly*, 23:54–58, March, 1952.

[47] Joseph E. Lantagne, "Health Interests of 10,000 Secondary School Pupils," *The Research Quarterly*, 25:330–333, October, 1952.

[48] James H. Humphrey, "Health Problems of Interest to College Men," *The Research Quarterly*, 23:150, October, 1952.

[49] Denver Public Schools, *Health Interests of Children*, Denver, Colo., Board of Education, 1947.

if they would like "to learn ways of preventing athlete's foot," "to visit a hospital and see an iron lung work," "to get the right amount of sleep," etc. Questions such as these usually appeal to youngsters and give a fair sample of interests at the moment.

The teacher at the local level in most communities can, with a little study, create a useful tool for appraising the interests of the students.

Health-Habit Questionnaire
(Revised)

Name ______ (Print last name first) Grade ______ Date ______

Instructions Please answer as carefully and accurately as you can each of the following questions concerning your health habits You are asked for this information in order that your physical education teacher may help you to improve your physical condition Your answers will be kept confidential

1 Are your living conditions congenial? Depressing? Do you have a room for yourself? Bed for yourself?

2 How many hours do you sleep each night? Is your sleep restful?

Do you sleep with your windows open at night? Are you warm at night (especially in the winter)?

3 Are you usually rested and refreshed in the morning? Drowsy?

Are you sleepy during the day? In class? When studying?

Do you take a nap during the day? How often? For how long?

Do you work and play without being more than comfortably tired mentally or physically at bed time? Fatigued?

Do you get to sleep easily at night? If not, why?

4 Are you often on edge nervous or jittery? Is it difficult for you to relax?

Are you subject to worries? Moods? Usually cheerful? Are you really happy?

5 How far do you live from school? How do you get to school?

What time do you leave in the morning? When home at night?

How much time do you usually study at home each school day?

How much time do you usually work at outside employment (or chores) each school day?

What do you do?

6 Do you have a hobby? What is it?

How many hours per day of physical activity do you usually get outside of school hours? What do you do?

What organizations do you belong to?

What social activities do you participate in with mixed groups (boys and girls)? How often?

What extra-curricular school activities do you take part in?

What other activities do you participate in?

7 Please check (X) the frequency with which you have the following

	Never	Seldom	Occasionally	Often
a. Headaches				
b Colds				
c. Sore throat				
d Ear ache				
e Indigestion				
f. Bad breath				
g. Coated tongue (bad taste)				
h. Pimples or skin eruptions				
i Boils				
j Twitching face and eyelids				
k. Eye strain				
l Sinus infections				
m Foot trouble				
n Joint pains				

Do you wear glasses? If so when were they last tested?

8. Do you eat three meals a day regularly? Is your appetite good? Do you eat at the school cafeteria at noon?

Carry your lunch? Go home for lunch? What do you usually eat at noon?

Do you eat between meals? (Check) Never Seldom Often Usually

What do you eat between meals?

9 How often do you usually eat each of the following kinds of food (check)

	Very Seldom	Once Each Week	Three Times Each Week	Once Each Day	Twice Each Day	Three Times Each Day
a Meat (including fish and eggs)						
b. Green vegetables (spinach, cabbage, lettuce, etc.)						
c. Other vegetables (carrots, peas, beans, beets, etc.)						
d Potatoes						
e Rice Macaroni						
f Pie, cake, pastry						
g Candy sweets						
h Fresh fruit						
i Salads						
j Oranges tomatoes						
k Dried fruits (prunes apricots figs etc)						
l Cereals						
m Pork						
n Fried foods						
o Whole wheat foods						

10 How many glasses of water do you usually drink daily? How many glasses of milk? Tea? Coffee?

11 Are you troubled with constipation? What do you do to correct it?

12 Do you smoke? If so how much daily?

Do you drink alcoholic beverages? If so what? How often? How much?

13 Do you visit a dentist once each year? Twice each year? How often do you usually clean your teeth each day?

14 Have you been vaccinated? Immunized for diphtheria? Typhoid?

What other immunizations?

15 Are your parents healthy and physically fit? If not what is the reason?

What is the physical stature of your father? { Tall Medium Short / Fat Average Thin

What is the physical stature of your mother? { Tall Medium Short / Fat Average Thin

16 Do you desire to be strong and physically fit? Do you wish to be attractive?

Are you satisfied with your present physical condition?

If your *Physical Fitness Index* is low can you account for it? How?

Summary of Interview

Safety Education Tests

Today, in the face of mounting accident rates, there is a strong effort made in most schools to stress safe living. Coupled with this is a tendency to treat the topic of safety as a separate item in the curriculum. It will always remain, however, part of the health education program.

Knowledge, attitudes, and habits relative to safety can be measured. This

is probably better accomplished in specific areas rather than in general. Thus in the elementary and secondary schools safety appraisal activities occur in special areas such as (1) safety in the home, (2) safety to and from school, (3) classroom safety, (4) school building safety, (5) playground safety, and (6) gymnasium safety. Additional special areas are (1) safety patrols, (2) bicycle safety, (3) traffic safety, (4) fire prevention and fire drills, and (5) first-aid procedures.

For most of the above topics there are few specific standardized tests available. There are on the community level, however, a number of administrative check lists and rating sheets that have been developed to help the teacher and administrator evaluate their safety efforts.

As already mentioned, there are several health knowledge and attitude measures that embrace material of a safety nature. These include tests such as the Crow-Ryan Health and Safety Education Test, the Shaw-Troyer Health Education Test, and the Mayshark Health and Safety Attitude Scale. In addition, there are a few more specific safety tests worth exploring for school use.

Myers Safety Attitude Scale for the Seventh Grade.[50] In this test the attitude of the student toward safety is measured by behavioristic questions (situation-response). Validity was established by a jury of experts. The reliability of the test is .89, a figure relatively high for a test of attitude measurement. Alternate Forms A and B are closely equivalent. Percentile norms have been established for both sexes in the seventh grade population. Here are two sample questions from the scale.

Form A

1. The sidewalk in front of your home is covered with ice. You would:

1. Put salt or ashes on the ice.
2. Tell your parents about the ice on the sidewalk.
3. Run and slide on the ice.
4. Be very careful when walking on it.

2. You see a teammate catching in a softball game without a head mask. You would:

1. Tell your teammate to wear a mask when he is catching.
2. Refuse to play ball until your teammate puts on a mask.
3. Throw a ball at his head to scare him.
4. Say nothing to your teammate.

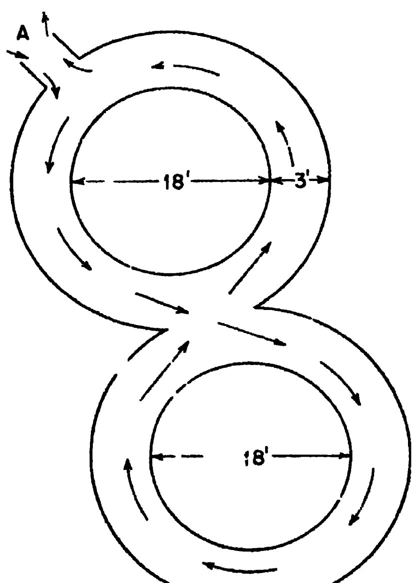
The National Safety Council has a number of useful tests at elementary and secondary levels for sampling pupil knowledge and understanding.[51]

[50] Frank H. Myers, "A Safety Attitude Scale for the Seventh Grade," *The Research Quarterly,* 29:320–325 October, 1958; address of author: Central Michigan College, Mount Pleasant, Mich.

[51] National Safety Council, *National Safety Council Test,* Chicago, Ill., The National Safety Council, 1946.

Test items have also been developed about fire prevention.[52] There are on occasion commercial sources of safety knowledge and skill tests for school use. One such test consists of twenty-five alternate-choice questions on bicycle safety, which are quite satisfactory with intermediate grade pupils.[53] Another test, a Bicycle Safety Quiz, is a twenty-question true-false test to be used in the classroom in conjunction with the motion picture, "Safe on Two Wheels," distributed free by the Public Education Department of Aetna Life Insurance Company.[54] This same company has an excellent teachers' guide on "How to Cycle in Safety." It also publishes a series of three bicycle skill tests. These can be set up on the school playground with a minimum of effort. Diagrams for laying out the testing area are very clear.

The following test is skill test 3, as used and recommended in "A Community Bicycle Safety Program."[55] The student begins to ride his bicycle at point A and rides slowly, being careful not to touch border lines.



Miscellaneous Techniques for Appraising Health Understanding and Behavior. A number of other measuring devices and techniques can be employed to evaluate student health knowledge, attitudes, habits, and interests.

[52] Hawkins, *National Fire Prevention Tests: General Achievement Tests in Fire Safety*, New York, New York University, Center for Safety Education, 1942.

[53] Association of Casualty and Surety Companies, 60 John St., New York 38, N.Y. (distributed free).

[54] Aetna Life Affiliated Companies, 151 Farmington Ave., Hartford 15, Conn.

[55] *A Community Bicycle Safety Program*, Association of Casualty and Surety Companies, 60 John St., New York 38, N.Y. (distributed free).

Check Lists and Rating Scales: Observation of health behavior of person or group is much more revealing if the instructor has something to use as a guide. The check list or rating scale brings order to observation. It is especially useful in noting day-to-day health attitudes and habits. To prepare a useful check list the teacher formulates a list of the more important outcomes expected for each unit or health topic. This can be reworked into a list of expected day-to-day practices and attitudes. The instructor's job is to fashion a simple check list of practical things that students do when they have a fairly good understanding of health practices. Moreover, these behavior signs can be rated numerically or by words to indicate the degree to which the student complies with the item. A desirable outcome of a teaching unit on personal appearance and cleanliness in the elementary grades is that the children assume responsibility for their personal appearance and that of others. On the teacher's check list under personal appearance and cleanliness might appear such questions as the following:

	Always	Usually	Seldom	Never
Are the hands, face, neck, and ears clean?				
Are the fingernails clean?				
Are the clothes neat and clean?				
Is the hair combed?				
Is a handkerchief or tissue carried?				
Are the hands washed before eating?				
Is the shower enjoyed following the physical education period?				
Is the mouth covered during coughing and sneezing?				
Is there good sitting and standing posture?				

Surveys: The survey is a formal approach to appraisal. It may be community-wide and involve information secured from parents and others, or it may be limited to the school. In either case it attempts to discover useful health information of use in the health teaching program. There may be a survey of the school cafeteria to check on the sale of milk, salads, fruits, whole-grain cereals, enriched bread, etc., or of a local restaurant, dairy, meat market, or grocery store. There are a variety of school surveys, but more important for health instruction is the survey that seeks to determine *student behavior* in such places as shower rooms and at such times as

relaxation periods. Environmentally, there may be an opportunity to wash hands before eating, but what does the survey show? How many students using a grade school lavatory actually wash their hands before eating? The survey, in this case, may make known a real need, which calls for a different teaching emphasis.

Closely related to the typical survey is the *pupil opinion survey* This is an especially useful device at the secondary level Here pupils are polled to ascertain personal and public health interests and to identify attitudes that may suggest school program changes Students like to express an opinion especially if they have had some hand in preparing the questionnaire or survey items

Student Diaries: Autobiographical material of boys and girls can provide a source of information about health status, health knowledge, and health attitudes that might not otherwise come to the attention of the teacher When considering the health problems of individual pupils, the diary is an appraisal device worth examining

Self-testing Activities: One cannot underestimate the significance of a student's own immediate problems and how *he* evaluates them It is not uncommon, therefore, to find personal health inventories and self-rating charts being used by the student to measure progress in terms of his own potential and to compare himself with his classmates It is desirable that a pupil sense that his feelings of adequacy, security, and acceptance are contingent on good grooming, skills, poise, proper health habits, and mental and emotional well-being Therefore, if the self-appraisal device is developed by the students who will use it and represents group decisions about items to be evaluated, it is likely to be far more meaningful than a standardized chart A pupil can readily rate himself on posture, skills, food habits, attitude toward others, personal appearance, and other such items Such self-testing activity will be successful if it is taken seriously and not to secure a grade or as a perfunctory class exercise

Oral Questioning: In small groups this is a fine method for understanding basic health principles and for concrete evidence of ability to apply those principles In fact, oral questioning can be more revealing than a standardized written health test It is time-consuming to question a whole school or class, but it can be used on a small sample basis where a random number of pupils are selected to express themselves

Health Conferences and Interviews: It would be most rewarding if teachers could sit down and talk at length with pupils about their personal health problems or to solicit views about school health practices. Most school personnel are too busy to do this, but it is well to keep in mind that it is an effective way to scrutinize the program Conferences and interviews need not be limited to pupils The school nurse, dental hygienist, physical director, or custodian are continually noting pupil expressions of behavior. Improper health habits and poor attitudes show up in every seg-

ment of the school curriculum and also permeate the extracurricular activities. The check list, rating scale, or survey form can often be used to guide the interview and keep it more to the point.

Parental Opinions: Frequently when boys and girls apply their knowledge, alert parents are the first to notice it. Attitudes and practices relative to visits to doctor and dentist, foods, personal cleanliness, oral hygiene, recreation, and rest are concrete items that sometimes cause parents concern; when students show increased cooperation along these lines at home, most parents will be aware of it and are happy to pass the information along to the teacher. In a small upstate New York town one teacher made it a point to talk to all twenty-eight mothers of her second-grade class. Using a simple check list, she asked pertinent questions about home improvement in specific areas where health instruction had been given After reviewing all the evaluation techniques that she had previously used, it was her judgment that the interview with the parents was by far the most fruitful and enlightening. As a specific example, it was found that after several lessons on activity, rest, and relaxation a rather large number of children willingly gave up television and late radio programs that might interfere with their sleep and went to bed.

Evaluation in Health Teaching

One of the chief purposes in testing for health knowledge and behavior is to improve the instructional program. Certainly the public is in need of help. Health misconceptions, superstitions, and gullibility appear in all walks of life. This is a nation of healthy persons, it is also a nation of tub thumpers, pill takers, addicts, quacks, faddists, and cultists Persons vary in their attitudes toward health on a scale all the way from sheer indifference, on the one hand, to the opposite state of overconcern leading to neurotic behavior. Proper evaluation at the school level can do much to reduce these weaknesses. Kilander, in his research on the public and its knowledge, has demonstrated that there is a high correlation between good health instruction in the schools and colleges and the scores obtained on his tests He has also shown that the relationship between health information and health practices is, in general, positive. A more thorough program, in which student progress in health instruction is regularly appraised, is one practice that can improve adult behavior.

SELECTED REFERENCES

Borozne, Joseph, and Leslie W. Irwin: "Prevalence of Certain Harmful Misconceptions among Prospective Elementary School Teachers," *The Research Quarterly*, 30: 460–468, December, 1960.

Branom, Wayne T.: "A Survey to Determine the Attitude, Habits, and Knowledge of Hillside, New Jersey, High School Students in Regard to Dental Health," *Journal of School Health*, 21:230–232, October, 1956.

Cushman, Wesley, and Bruce Bennett: "Checklist of Health Problems," *Journal of Health, Physical Education and Recreation,* 30:28–29, December, 1959.

Dearborn, Terry H.: "Personal Health Knowledge of College Students before Instruction," *The Research Quarterly,* 29:154–159, May, 1958.

Irwin, Leslie W., Charles D. Merrill, and Wesley M. Staton: "Concepts of Healthful Living of Fundamental Value in the General Education of Elementary School Pupils," *The Research Quarterly,* 24:422–415, December, 1953.

Patty, Willard W.: "Reading Difficulty Differences of Health Knowledge Tests," *The Research Quarterly,* 16:199–201, October, 1945.

———: "Blanket Health Tests Unsatisfactory," *Journal of School Health,* 20:258–260, November, 1950.

Richardson, Charles E.: "Sentence Completion Health Attitude Test for College Students," *Journal of School Health,* 30:33–35, January, 1960

Turner, Claire E., C. Morley Sellery, and Sara Louise Smith: *School Health and Health Education,* St. Louis, The C. V. Mosby Company, 1957, chap. 20.

Washnik, Walter M.: "What the High School Senior Knows about Health," *The Research Quarterly,* 28:178–179, May, 1957.

Willgoose, Carl E.: *Health Education in the Elementary School,* Philadelphia. W. B. Saunders Company, 1959, chap. 12.

CHAPTER 6

Measures of Physical Fitness Cardiovascular Tests

If you can force your heart and nerve and sinew
To serve your time long after they are gone
And so hold on when there is nothing in you
Except the Will which says to them: "Hold on!"

Rudyard Kipling

The noblest thoughts in the minds of men are but wishful thinking in a body physically unable to put the thoughts in action. Kipling's "Hold on" requires physical strength. The virtuous behavior of mankind and the rewards of a rich and full life are tied closely to man's physical capacity.[1] Physical strength to act and muscular and cardiorespiratory endurance are still essential in the structure of a civilization.

The Need for Physical Fitness Measurement

More persons in present-day society have an interest in physical fitness than at any other time since 1900 Medical research, especially in cardiology, physical medicine, and psychiatry, has indicated the importance of physical activity in maintaining and promoting optimum health.

School children need a minimum level of physical fitness. It is here particularly that physical education personnel make a most essential and unique contribution.[2] The relationship between health status, health behavior, and physical performance suggests that more attention be given to an analysis of motor activity Although physical fitness tests are not designed to measure such health items as nutritional status, chronic disease, or emotional disturbances, they nevertheless frequently reflect weaknesses in these areas. Because of this factor alone they are valuable as screening devices whereby children and adults who are substrength and lack the capacity

[1] The ability of the organism to do, to be able, to sustain effort, to resist fatigue and recover rapidly, and to perform tasks with sufficient resources left to meet an emergency.

[2] See article defending this point of view by Carl E. Willgoose, "Physical Fitness—The Primary Objective," *Journal of Health, Physical Education and Recreation,* 30:33–36, November, 1959.

to perform physically can be detected early and examined more thoroughly to determine specific organic strains and drains. A school health program, therefore, that fails to test for physical fitness is incomplete.

Definition of Terms

There is an interrelationship of terms commonly used in physical education. Physical fitness, motor fitness, general motor ability, and cardiovascular endurance have much in common, but they are terms that should not be used synonymously.

Physical fitness is a capacity for sustained physical activity. It can be demonstrated through motor activity involving the total organism. More specifically, it is *quantitative* and can be evaluated by measuring a person's muscular strength, muscular endurance, and cardiovascular endurance. Although very much related to total health or total fitness, it is nevertheless a distinct and tangible element.

Motor fitness is frequently used synonymously with physical fitness because it involves physical performance items that show muscular strength and endurance. Motor fitness is more than this. It is a *qualitative* performance involving some degree of personal skill. It consists of power, flexibility, speed, agility, and balance. For example, a test of leg strength would give a measure of physical fitness of the legs because a quantity of strength and endurance is measured. On the other hand, a more elaborate test would measure the quality and functional efficiency of the legs by testing speed, agility, and power. The purpose of this illustration is to indicate that motor fitness is more general than physical fitness. Both are founded on sound nutritional status and organic condition. Both are health measures useful in screening school children.

General motor ability is coordination of the various body parts in movement. It involves proprioceptive (kinesthetic) arm-eye-foot coordination demonstrated through physical skills in handling the body. It is related to physical fitness in that the performance of motor skills rises and falls according to the elements of muscular strength and endurance.

Cardiovascular endurance, also referred to as cardiorespiratory endurance and circulatory endurance, is a kind of physiological fitness demonstrated through an adjustment of the heart and lungs to prolonged physical exertion.

Muscular strength is the ability of a muscle or muscle group to exert maximum strength in a single contraction.

Muscular endurance is the ability of a muscle or muscle group to maintain a submaximal contraction over a period of time.

For purposes of organization this topic is divided into three parts. This chapter is concerned with cardiorespiatory tests. Chapter 7 is about tests of muscular strength and muscular endurance. Chapter 8 concerns numerous tests of fitness involving a more elaborate physical performance than is involved in measuring pure strength and endurance. These tests require such

activities as chinning, jumping, and running against time. There is considerable overlapping, however, in the use of the terms *physical* and *motor* when speaking of fitness tests. In fact, most tests labeled motor fitness tests are generally thought of as measures of physical fitness. Therefore, one will observe that tests such as the Indiana Motor Fitness Test, the Oregon Motor Fitness Test, the California Physical Performance Test, and the New York State Physical Fitness Test are in the same category.

Cardiovascular Tests of Fitness

As more and more physical work is performed, the efficiency of the muscular system improves. With the increase in muscle tone, power, and endurance, an increase in functional ability is reflected in all organic systems of the body. Increased efficiency occurs especially in the heart and lungs. There is a fuller exchange of oxygen and carbon dioxide both in the lungs and in the muscles. With muscle training the stroke volume of the heart improves to a point where more blood is pumped per stroke, with fewer strokes per minute than in the untrained or unconditioned person. In short, the total cardiac output is increased. Furthermore, the rate of lactic acid formation is lowered in the trained person, resulting in a lower blood-lactate concentration. There is also a slower rate of breathing and a corresponding economy in respiration.

Relationship of Muscular Strength and Endurance to Cardiovascular Endurance

Cardiovascular or cardiorespiratory measures usually attempt to appraise heart-lung efficiency during or after some specified amount of physical exercise. Cardiovascular endurance, although a separate variable in physical fitness, is very much related to muscular endurance. When a muscle is involved in an endurance activity such as floor push-ups or running the mile, the circulatory and respiratory systems work harmoniously to remove waste products quickly. Measures of cardiovascular endurance do not always correlate highly with those of muscular endurance. This is not to say that they do not rise and fall together, but the circulatory-respiratory function seems to be affected by more external stimuli (digestion, age, climate, emotional state, attitude of mind, etc.) than does a particular kind of muscular endurance.[8] Also, in many cases after prolonged training, cardiovascular endurance may remain high without much change, whereas further physical conditioning tends to increase muscular endurance.

Because the efficiency of most internal organs of the body influences the heart and lungs, cardiovascular measures are frequently considered as screening tests of general organic efficiency.

The statistical method of factorial analysis is being used more and more to determine the basic components of different tests of physical fitness and

[8] Leonard A. Larson, "Cardiovascular-Respiratory Function," *The Research Quarterly* (Supplement), 12:456–459, May, 1941.

other mental and athletic abilities In analyzing tests of endurance an effort is made to ascertain the "pure" factors that are unrelated each to the other McCloy, working with a series of twelve athletic events administered to 400 well-conditioned soldiers, found that four distinct factors were present.[4] They were circulorespiratory endurance, velocity, muscular endurance, and mesomorphic body build. McCloy's tests of cardiovascular endurance were performance items with considerable heart-lung effort over a period of time They included a combination of the 300-yard run, a 440-yard run, a mile run, a sixty-second squat-thrust, and a 75-yard run with a man astride the back McCloy postulated that these tests were in themselves highly potent measures of cardiovascular endurance.

Exercise of long enough duration to reach the endurance stage is generally useful in showing a rather strong relationship between muscular endurance and cardiovascular endurance This was demonstrated by McCurdy and Larson when they were validating their test of organic efficiency[5] Working with infirmary patients in poor physical condition and varsity college swimmers, they were able to show correlations of 68 and 70 between their organic efficiency test items and performance on the 440-yard swimming event

Although positive correlations usually occur between tests of cardiovascular endurance and muscular endurance, they are generally too low to be useful. They indicate that there are definitely two distinct endurance variables, and these variables appear to correlate with one another more when an all-out sustaining physical exercise is performed The short burst of speed, for example, or the completion of fifteen floor-dips does not greatly disturb the heart-lung apparatus compared with swimming or running the quarter mile The same reasoning applies to pure muscular strength Unless the activity is repeated sufficiently to activate the cardiorespiratory apparatus, little change in heart-lung endurance will be noticed This is especially true in the case of well-conditioned men and women Cureton obtained low correlations varying between 002 and .31 with over two dozen tests of strength and endurance when compared with the results of the Harvard Step Test, a test of organic efficiency[6] Bookwalter obtained similar results after comparing the scores on the same test with the scores on the Army Physical Fitness Test, which he administered to 1,269 cadets in the Army Specialized Training Program[7] Although Taddonio and Karpovich obtained a higher rank-difference correlation between the short form of the same cardiovascular test and intramural cross-country running (63), the score

[4] Charles H McCloy, "A Factor Analysis of Tests of Endurance," *The Research Quarterly*, 27:213–216, May 1956

[5] James H McCurdy and Leonard A Larson, "Measurement of Organic Efficiency for the Prediction of Physical Condition," *The Research Quarterly* (Supplement), 6:9–14, May, 1935

[6] Thomas K Cureton, *Physical Fitness Appraisal and Guidance*, St. Louis, The C V. Mosby Company, 1947, p. 288

[7] Karl W. Bookwalter, "A Study of the Brouha Step Test," *The Physical Educator*, 5:8, May, 1948.

was still too low to be significant.[8] A survey by Clarke of the relationship between strength tests and cardiovascular tests yielded correlation coefficients varying from —.04 to .55; approximately three-fourths of the correlation coefficients fell below .20.[9]

Breath Holding as a Measure of Cardiovascular Condition

Although breath holding has long been recognized as related to physical condition, the particular relationship has not been significant. This has limited the use of breath-holding tests.

There is undoubtedly a decreased demand for oxygen and a corresponding decrease in the respiratory rate of the trained person. This has been shown over the years by a number of investigations For example, Henry demonstrated that twenty-one subjects with severe colds dropped 18.1 per cent in the flarimeter breath-holding test after ninety seconds of 12-inch stool stepping at thirty steps per minute[10] Jokl found significant increase in breath holding accompanied by physical training.[11] In six months, thirty-two experimental subjects practically doubled their breath-holding time. There are difficulties, however, in putting this research to use. There are at least two reasons why these tests are of questionable value:

1. Breath holding is partly influenced by such psychological factors as will power and motivation

2. Breath holding tends to improve with practice by partial strengthening of the expiratory muscles This may be out of proportion to any real improvement in the circulorespiratory apparatus and body as a whole.

PULSE RATE MEASUREMENT

The pulse rate can be measured in a number of ways Either the carotid or radial pulse can be used To measure the former, three fingers are placed on the carotid artery just below the ear, at the base of the neck. To measure the radial pulse three fingers are placed over the radial artery, which lies in a groove on the thumb side of the wrist. Frequently the count is recorded for thirty seconds. The subject must be prepared in accordance with specific test instructions A resting pulse, (that is, a pulse rate representative of a heart-lung system at ease) is frequently a requirement in pulse rate tests.

BLOOD PRESSURE MEASUREMENT

A number of cardiovascular measures require an appraisal of blood pressure. This is accomplished with the mercury sphygmomanometer and a

[8] Dominick A. Taddonio and Peter V. Karpovich, "The Harvard Step Test as a Measure of Endurance in Running," *The Research Quarterly*, 22·381–384, October, 1951.

[9] H. Harrison Clarke, *The Application of Measurement to Health and Physical Education*, Englewood Cliffs, N.J., Prentice-Hall, Inc., 1959, p 113.

[10] Thomas K. Cureton, *op. cit.*, p. 349.

[11] E. Jokl et al., *Training and Efficiency*, Johannesburg, South Africa, The South African Institute for Medical Research, 1941, p. 84.

stethoscope. The arm is bared, relaxed, and slightly flexed. It rests on a smooth surface at about the level of the heart. The cuff is wrapped snugly around the arm just above the joint of the elbow. The receiver of the stethoscope is then placed firmly over the brachial artery in the antecubital space. The cuff is pumped up to a point where no pulse beat can be heard. Pressure is then slowly released by the tester. As this is done, a close watch is kept on the mercury column. At the first sound of pulse the mercury column is read. This reading is in millimeters of mercury and is referred to as the *systolic* pressure. As the tester continues to let pressure out of the cuff, the pulse beat will change from a strong beat to a very weak beat. At this point the column of mercury is read again. This reading is known as the *diastolic* pressure. In recording blood pressure readings the systolic reading is placed over the diastolic reading. For example, a measurement of 120/80 signifies a systolic pressure of 120 millimeters of mercury over a diastolic pressure of 80 millimeters of mercury.

BARACH INDEX

One of the cardiovascular tests in use for many years is the Barach Energy Index, which determines the amount of energy the heart expends in blood output.[12] The test employs systolic and diastolic blood pressure together with the pulse rate count per minute. The subject rests before taking the test so that a resting pulse count can be obtained. The test is given to the subject in a sitting position. It is computed as follows:

$$\text{Energy index} = \frac{\text{systolic pressure} + \text{diastolic pressure} \times \text{pulse rate}}{100}$$

Thus, a person with a 120/80 blood pressure measurement and a pulse rate count of 60 would have an index of 120. It has been shown that the index varies between 110 and 160 for generally healthy persons, with upper and lower score limits of 200 and 90, respectively. Scores below 90 indicate hypotension, and scores above 200 indicate hypertension.

In studying 200 college men at the University of Illinois, Cureton obtained an average Barach index of 140.85 in a range of scores from 70 to 220.[13] Hunsicker also worked with college students. He found that this test correlated fairly well with a measure of cardiac output, namely, the heart-stroke volume divided by body surface area for an all-out treadmill run.[14] The coefficient of correlation was —.50, indicating to some degree that the more favorable the cardiac output the lower the Barach index score.

[12] J. H. Barach, "The Energy Index," *Journal of the American Medical Association,* 62:525–530, Feb. 14, 1914.

[13] Thomas K. Cureton, *op. cit.*, p. 285.

[14] Paul A. Hunsicker, "A Validation of Cardiovascular Tests by Cardiac Output Measurements," unpublished doctoral dissertation, Univertiy of Illinois, Champaign-Urbana, Ill., 1950.

FOSTER'S TEST

The basic premise in Foster's test is that the heart rate increases in proportion to physical exercise Also, during the rest period following the exercise, the heart rate returns to the beginning rate at a speed consistent with the physical condition of the subject.[15]

In this test the pulse rate per minute is recorded in the standing position The subject exercises for thirty seconds by running in place at a rate of 180 steps per minute. Immediately following this, the pulse rate is taken for five seconds and multiplied by twelve to give a rate per minute The subject continues to stand at ease for forty-five seconds, at the end of which his pulse rate is recorded again. Points are awarded by Foster for the level of standing pulse rate and the level after exercise and after rest Foster's research indicates that among boys in the age group fourteen to nineteen there would not be an increase greater than forty beats per minute for those in good physical condition Furthermore, a rest of forty-five seconds should be sufficient to permit the person in satisfactory condition nearly to recover from the exercise

The test is weak in reliability With a more standardized and controlled procedure, however, it could have some value as a quick means of screening As it is, the principles involved are applied routinely in many a physician's office to obtain a rough idea of general physical condition

HEARTOMETER

After considerable research, Cureton pointed out that the Cameron heartometer can be used in health and physical education as a screening device to "show differences between normal individuals in present cardiovascular condition and relative amounts of cardiovascular fatigue"[16] The heartometer provides a heartograph (or heartogram), a graphic record made by a pen activated by the pulsations of the brachial artery transmitted by blood pressure apparatus of the sphygmomanometer type The writing pen is activated by air pressure and leverage The energy of the pulse wave is shown vertically, and time lapse is shown horizontally The heartometer differs from the electrocardiograph in that it registers variations of mechanical pressure rather than electrical variations It was validated against endurance items that included distance running, distance swimming, treadmill running, and step-test performance.[17]

Although this test seems to be a valid measure for appraising cardiovascular efficiency and norm tables for sex and age are available, it appears,

[15] W. L. Foster, "A Test of Physical Efficiency," *American Physical Education Review*, 19:20–24, December, 1914.

[16] Thomas K. Cureton, *op. cit.*, p 232.

[17] Thomas K. Cureton et al., *Endurance of Young Men*, Society for Research in Child Development, National Research Council, Washington, 1945.

nevertheless, to be far too complicated a measure to use in schools. Considerable familiarity with the heartometer and two other related instruments is necessary (the Esser Planimeter and the Vernier calipers). Moreover, experience is needed to analyze and interpret the test scores.

CRAMPTON'S BLOOD PTOSIS TEST

Shortly after the turn of the century much of the cardiovascular research was conducted by C. Ward Crampton. Crampton's objective was to find a test of general physical condition that could be used in growth and development studies of youth. This research led to one of the earliest known heart-lung efficiency measures.[18] It is based on heart-rate changes and systolic blood pressure upon standing up after having been in a reclining position.[19] During such an action, according to Crampton, heart rate rises from zero to forty-four beats per minute, and systolic blood pressure varies from —10 to +10 millimeters of mercury Furthermore, in persons in good physical condition systolic pressure rises 8 to 10 millimeters of mercury when the erect position is taken, whereas in the poorly conditioned person pressure may not rise at all; in fact, it may drop. Moreover, the pulse rate in the poorly conditioned person may rise considerably after he stands, whereas the pulse rate of the physically fit person may not rise at all.

Working with secondary school students in New York City, Crampton assigned equal values to these systolic blood pressure and heart-rate variations. This vascular tone is expressed as shown in the following table:

Systolic Blood Pressure

Heart-rate increase	Increase					0	Decrease				
	+10	+8	+6	+4	+2		−2	−4	-6	−8	−10
0 4	100	95	90	85	8C	75	70	65	60	55	50
5–8	95	90	85	80	75	70	65	60	55	50	45
9-12	90	85	80	75	70	65	60	55	50	45	40
13–16	85	80	75	70	65	60	55	50	45	40	35
17 20	80	75	70	65	60	55	50	45	40	35	30
21 24	75	70	65	60	55	50	45	40	35	30	25
25 28	70	65	60	55	50	45	40	35	30	25	20
29–32	65	60	55	50	45	4)	35	30	25	20	15
33–36	60	55	50	45	40	35	30	25	20	15	10
37–40	55	50	45	40	35	30	25	20	15	10	5
41–44	50	45	40	35	30	25	20	15	10	5	0

SOURCE: C. Ward Crampton, "A Test of Condition: Preliminary Report," *Medical News*, 87:529-530, September, 1905.

[18] C. Ward Crampton, "A Test of Condition: Preliminary Report," *Medical News*, 87:529–530, September, 1905.
[19] C. Ward Crampton, "Blood Ptosis," *New York Medical Journal*, Nov. 8, 1913.

Test Equipment Needed: A watch with a second hand, blood pressure measurement apparatus, a bench to recline on, and a low pillow for the head are needed.

Testing Procedure: 1. Place the subject in a comfortable reclining position with pillow under the head. He should stay in this position until the pulse rate is constant. When two repeated fifteen-second counts are the same, the pulse rate may be considered constant (resting pulse).

2. Record the pulse rate for one minute.

3. Record the systolic blood pressure.

4. Have the subject stand. Record the standing pulse rate; to be certain that the standing pulse rate is accurate, recheck the count so that two fifteen-second counts are the same.

5. Record the systolic blood pressure in the standing position.

Scoring: Compute the difference between (1) the reclining and standing pulse rates and (2) the reclining and standing systolic blood pressure readings. Then consult the above table to secure the test score.

Example

Pulse rate		*Systolic blood pressure*	
Reclining	70	Reclining	100
Standing	85	Standing	110
Difference	+15	Difference	+ 10

According to the chart, a pulse rate difference of +15 and a systolic pressure difference of +10 yields a test score of 85.

Interpretation: Crampton considered a test score in the range 60–100 representative of good health. A person scoring below 50 should have a special examination to ascertain the reason. A score below zero would indicate a very severe state of disturbance in the circulatory system.

By examining several hypothetical cases with the scoring chart it is possible to obtain test scores covering a wide variation in heart rate and blood pressure readings, all within the normal range. Thus the Crampton test seems to be a more useful measure for evaluating hosiptal patients in an extremely low state of physical fitness. McCloy shares this viewpoint and has indicated that the test fails to differentiate between persons in relatively good physical condition.[20]

CARLSON FATIGUE CURVE TEST

Because of the difficulty in measuring cardiovascular differences between persons in relatively good physical condition, Carlson designed a test to

[20] Charles H. McCloy and Norma D. Young, *Tests and Measurements in Health and Physical Education*, New York, Appleton-Century-Crofts, Inc., 1954, p. 291.

force the subject to work hard and thereby demonstrate an all-out example of physical condition[21] This kind of test, therefore, would have value in physical education In fact, Matthews points out that it is sufficiently precise to differentiate objectively students in physical education classes.[22]

In this test the subject runs as fast as he can in place for ten seconds It is an all-out effort, following which he rests for ten seconds This is repeated ten times The pulse is taken before exercise, ten seconds after exercise, and two, four, and six minutes after exercise Also, a count is made of the number of times the right foot hits the floor during a ten-second period of exercise The total number of floor contacts for the ten periods is known as *production* In running, the subject lifts his feet just far enough to clear the floor.

Test Equipment Needed A watch with second hand, paper, and pencil are needed

Testing Procedure 1 Seat subject on gymnasium floor or ground Record the pulse rate (resting pulse)

2 Have the subject perform the ten-second run in place Follow this by a ten-second rest in place

3 Repeat nine more times Encourage subject to do his very best throughout the testing activity Poor motivation will invalidate the test score.

4 Record the pulse rate ten seconds after the final period of exercise

5 Record the pulse rate two minutes after exercise

6 Record the pulse rate four minutes after exercise

7 Record the pulse rate six minutes after exercise

Scoring The ten period scores can be kept on a single piece of paper with the five different pulse rate scores Students can keep their own scores, thus enabling a teacher to test a whole class in a ten-minute period Each set of scores can be plotted on graph paper as follows

1 Plot a fatigue curve showing *production* (total number of right-foot contacts) on the ordinate and periods on the abscissa

2 Plot a fatigue curve showing the *total pulse-rate count* on the ordinate and time lapse of pulse count on the abscissa (that is, before exercise, after exercise, two minutes later, four minutes later, and six minutes later)

3 Have each student keep his own graphs so that future test scores can be plotted on the same graphs

Interpretation Repeated tests bouts in short period of time serve as a conditioning activity and tend to improve succeeding performances There is some indication that students enjoy this kind of check on personal physical condition and that it helps to screen those persons who are at the low end of the scale This measure might have more value if norm tables were available for age and sex.

[21] H C Carlson, "Fatigue Curve Test," *Research Digest*, 16 72–78, October, 1945.
[22] Donald K Matthews, *Measurement in Physical Education*, Philadelphia, W. B. Saunders Company, 1958, p 190

McCURDY-LARSON ORGANIC EFFICIENCY TEST

In an effort to find a test for college men that would estimate the functional efficiency of the circulatory-respiratory systems, McCurdy and Larson experimented with over two dozen measures and selected five for their test battery.[23] The five items are (1) sitting diastolic blood pressure, (2) breath holding twenty seconds after exercise, (3) the difference between standing-resting pulse rate and standing pulse rate two minutes after exercise, (4) standing pulse pressure, and (5) vital capacity. Age and weight are also figured.

McCurdy and Larson validated their test on the performance of varsity swimmers as the "good" group and infirmary patients in a febrile state after respiratory infection as the "poor" group. Test validity was .83, and the multiple correlation of the five test items was .947. Norms have been prepared by Dane for boys in the age range 10½ to 17½ years.[24] These norms are divided into three physiological age groups: prepubescent boys, pubescent boys, and postpubescent boys.

Test Equipment Needed: The equipment needed includes weight scales, blood pressure apparatus, stop watch, and a flarimeter.[25] A set of two-step exercise stairs is also needed. In order to complete the test on one subject in fifteen minutes, it is suggested that six assistants at six testing stations be employed to obtain the raw data.

Testing Procedure: In this test the subject is weighed. Sitting diastolic, standing systolic, and standing diastolic blood pressure measurements are made. Standing pulse rate is recorded. Vital capacity is recorded, and breath-holding practice is permitted on the flarimeter. Then a stair-climbing exercise is performed on the steps for ninety seconds at a predetermined rate. This is followed by breath holding twenty seconds after exercise, and the standing pulse rate is taken two minutes after exercise.

Scoring: Raw scores are transformed into weighted standard scores. Standard scores are added together to provide an index for classification.

Interpretation: At a later date Larson shortened his test to three items: sitting diastolic pressure, breath holding after exercise, and standing pulse pressure.[26] He also developed a Cardio-Respiratory Rating Scale, which can

[23] James H. McCurdy and Leonard A. Larson, "Measurement of Organic Efficiency ...e Prediction of Physical Condition," *The Research Quarterly* (Supplement), ...41, May, 1935; see also "The Measurement of Organic Efficiency for the Pre...tion of Physical Condition in Convalescent Patients," *The Research Quarterly,* 78–97, December, 1935.

[24] C. Wesley Dane, "A Study of Circulatory-Respiratory Changes as Indicated by the McCurdy-Larson Organic Efficiency Test in Relation to Physiological Age," *The Research Quarterly,* 15:98–112, May, 1944.

[25] An instrument employed to measure both breath-holding capacity and vital capacity.

[26] Leonard A. Larson, "A Study of the Validity of Some Cardiovascular Tests," *Journal of Experimental Education,* 3:214–218, March, 1939.

be used with single test items and scored excellent, very good, good, fair, poor, and very poor

The test is complicated In using it one must refer to several tables, read instructions to the subject, and spend at least fifteen minutes with each person being tested Furthermore, the exercise period for a predetermined number of steps on the special step-up stairs is difficult to standardize Also, the flarimeter seems to be of low validity as a measure of respiratory efficiency.

Breath holding is partly a matter of will power, and this seems to improve with practice against known standards [37] Although the test may have some value with hospital convalescents it does not seem to measure critically physical fitness items such as endurance in running and swimming Olds found this to be true [38] So did Cureton in a study showing low correlations between test data and an experimental group of normal subjects and well-conditioned swimmers [39] In all fairness to the authors of the test it should be pointed out that the test works better with subjects returning to normal after a period of temporary illness As Weiss and Phillips warn, it is a limited test to note "improvement in physiological fitness" and does not predict success in athletic events or measure motor ability [40]

SCHNEIDER TEST

The test of E C Schneider is one cardiovascular test to stand up over the years as a valid measure of physical condition [41] The six test items were arbitrarily established and tried out on 2,000 aviators during World War I

Test Equipment Needed Watch with second hand, blood pressure apparatus, cot for reclining, and chair to step up on (18½ inches high) are needed

Testing Procedure 1 Have the subject being tested recline on the cot for five minutes

2 While the subject is still reclining take the pulse rate for twenty seconds Repeat at twenty-second intervals until two consecutive counts are the same Multiply this count by three to convert to pulse rate per minute Record the score

3 While the subject is reclining record the systolic blood pressure This may be repeated to obtain accuracy

4 Have the subject stand for two minutes

5 Measure pulse rate for fifteen seconds Repeat at fifteen-second

[37] Thomas K Cureton, *Physical Fitness Appraisal and Guidance*, p 350

[38] L W Olds 'Study of the Effect of Competitive Basketball upon the Physical Fitness of High School Boys as Determined by the McCurdy-Larson Organic Efficiency Tests" *The Research Quarterly* 12 254–265 May 1941

[39] Thomas K Cureton, *Physical Fitness Appraisal and Guidance*, p 299

[40] Raymond A Weiss and Marjorie Phillips, *Administration of Tests in Physical Education* St Louis, The C V Mosby Company 1954 p 61

[41] E C Schneider 'A Cardiovascular Rating as a Measure of Physical Fatigue and Efficiency," *Journal of the American Medical Association*, 74 507, May 29, 1920.

intervals until two consecutive counts are the same. Multiply this count by four to convert to pulse rate per minute. Record the score.

6. Record the difference between the standing and reclining pulse rate.

7. With the subject standing, record the systolic blood pressure.

8. Record the difference between the standing and reclining systolic blood pressure.

9. Use a stop watch. Have the subject step up on a chair 18½ inches high five times in fifteen seconds (once every three seconds).

10. Record the pulse rate for fifteen seconds immediately after the exercise. Multiply by four to obtain the rate per minute.

11. With subject still standing, continue taking pulse rate at fifteen-second intervals until the subject's pulse rate has returned to normal. Record the number of seconds it takes for the return to normal. This is accomplished by comparing the pulse rate after exercise with the first normal fifteen-second count. Where the rate is not back to normal in a two-minute period, record the number of beats still above normal.

Scoring. By referring to Table A-23, Appendix A, the pulse rate and systolic pressure changes (standing, after exercise, and in recovery) can be scored from +3 to −3. By combining all six test scores a maximum or perfect score of 18 may be obtained. Anything less than a total of 9 is considered a deficiency and requires further medical examination.

Interpretation. The Schneider test is a mild test of circulatory efficiency, which will discriminate cardiovascular condition only in gross categories when persons of widely differing health status are measured. As such, it has value as a screening device for convalescent persons and others of extremely poor physical condition.

Although it has been used in a great many research projects to note heart-lung efficiency, there is still some question about its value as an appraisal instrument in physical education circles. Certainly the reliability factor is an important consideration. Taylor and Brown raise this question.[32] So does Cureton, but he obtained a .89 reliability coefficient.[33] When compared with physical performance, the Schneider test appears weak. Henry and Herbig compared the test with improvement time on the half-mile run and obtained a coefficient of correlation of .41.[34] McCloy and Young had a similar correlation coefficient when comparing test scores with health status.[35] In short, the test measures a unique fitness element, heart-lung condition, and does not adequately relate it to strength, skill, or power.

[32] C. Taylor and G. E. Brown, "Some Observations on the Validity of the Schneider Test," *Journal of Aviation Medicine*, 15: 214-330, 1944.

[33] Thomas K. Cureton et al., *Endurance of Young Men*, p. 214.

[34] Franklin Henry and W. Herbig, "The Correlation of Various Functional Tests of the Cardio-circulatory System with Changes of Athletic Condition of Distance Runners," mimeographed report presented to Research Section, American Association for Health, Physical Education, and Recreation, San Francisco, Calif., 1939.

[35] Charles H. McCloy and Norma Young, *op. cit.*, p. 292.

HARVARD STEP TEST

Another measure of cardiovascular condition is the five-minute step test developed during World War II at the Harvard University Fatigue Laboratory.[36] It is designed to measure fitness for muscular work and ability to recover. Brouha tested 2,200 male college students for endurance in treadmill running, blood-lactate level, and maximum heart rate per minute. He found that athletes outperformed nonathletes and increased their scores with additional training.[37] Later Taddonio and Karpovich supported the relationship with physical performance by obtaining a correlation of .63 between the step test and the order in which college men finished a cross-country race.[38]

Test Equipment Needed: A watch with a second hand and a bench or platform 20 inches high are needed.

Testing Procedure: This can be an individual or a group test. In the latter a leader counts cadence for the whole group of subjects. Each subject has an observer who measures the pulse counts. A new group can be tested every ten minutes, the size of which depends upon the number of benches available. Prospective subjects, especially in a class situation, should remain seated and as quiet as possible before being tested. As in other cardiovascular measures, a resting pulse rate is desired to begin with. In an individual testing situation, proceed as follows:

1. Have the subject stand up and face the step-up bench. Particular attention should be paid here to indicate the method and rhythm used in stepping. Count as follows: left foot *up;* right foot *up;* left foot *down;* etc. This is a four-count rhythm. Subjects should breathe fully throughout and straighten the knees completely on top of the bench. With the help of a watch the cadence can be kept to thirty steps per minute.
2. Begin to count by saying "ready,' "go." Continue counting for five minutes unless the subject quits from exhaustion before the end of that period.
3. At the end of five minutes give the command to "stop" and have the subject sit down.
4. Exactly one minute later record the pulse rate for thirty seconds. Exactly thirty seconds later measure and record the pulse rate for another thirty seconds. Finally, thirty seconds later take the third recovery pulse rate for thirty seconds, and record. Thus, you will have taken the recovery pulse count from 1 to 1½, 2 to 2½, and 3 to 3½ minutes after the exercise is terminated.

[36] Lucien Brouha, "The Step-Test: A Simple Method of Measuring Physical Fitness for Muscular Work in Young Men," *The Research Quarterly,* 14:31–36, March, 1943.

[37] Lucien Brouha, Norman W. Fradd, and Beatrice M. Savage, "Studies in Physical Efficiency of College Students," *The Research Quarterly,* 15:211–244, October, 1944.

[38] Dominick A. Taddonio and Peter V. Karpovich, *op. cit.,* p. 381.

5. The above figures may be conveniently recorded on a card somewhat as follows:

Name of subject	Date	Duration of exercise (seconds)	*Recovery counts after exercise (minutes)* 1	2	3	Total recovery count	Index of PE

Scoring: Introduce into the formula the duration of exercise in seconds and the total of the three half-minute pulse counts. This will yield the index of physical efficiency.

$$\text{Index of PE} = \frac{100 \times \text{duration of exercise (seconds)}}{2 \times \text{total pulse counts in recovery}}$$

Compare the index of PE with general physical condition terms as follows:

Physical condition	*Index of PE*
Excellent	90 and above
Good	80-89
Average	65 79
Low average	55-64
Poor	54 and below

Test Modifications: A *short form* Harvard Step Test was developed by Johnson and Robinson in the interest of time.[39] It is based on the duration of exercise and a single pulse rate count taken 1 to 1½-minutes after exercise. Data are introduced into the formula as follows:

$$\text{Index of PE} = \frac{100 \times \text{duration of exercise (seconds)}}{5.5 \times \text{pulse count 1 to } 1\frac{1}{2} \text{ minutes after exercise}}$$

This index of physical efficiency is interpreted as follows:

Above 80 = Good physical fitness
50–80 = Average physical fitness
Below 50 = Poor physical fitness

Other Harvard Step Test modifications have been made in an effort to consider age of subject and body size. Gallagher and Brouha developed a modification for secondary school boys, aged twelve to eighteen.[40] The bench was lowered to 18 inches, the duration of exercise was reduced to four minutes, and body surface was calculated from height and weight measurements. See Table A-6, Appendix A for nomographic chart used for computing the surface area. A boy with a surface area greater than 1.85

[39] See Edward C. Schneider and Peter V. Karpovich, *Physiology of Muscular Activity*, Philadelphia, W. B Saunders Company, 1953, p. 270.

[40] J. R. Gallagher and Lucien Brouha, "A Simple Method of Testing the Physical Fitness of Boys," *The Research Quarterly*, 14:23–30, March, 1943.

square meters would use a 20-inch bench. The scoring is the same as for college men. When the regular index of physical efficiency formula is applied, a secondary school boy's score would be interpreted as follows:

Above 90	= Superior physical condition
81-90	= Excellent physical condition
71-80	= Good physical condition
61-70	= Fair physical condition
51-60	= Poor physical condition
Below 51	= Very poor physical condition

In this test the authors found that approximately 50 per cent of the cases fell at the level of fair physical condition.

Gallagher and Brouha also worked out a variation of the Harvard Step Test for *high school girls*.[41] The bench height was reduced to 16 inches and the exercise duration to four minutes. The scoring procedure was not changed, but instead of referring to an index of physical efficiency they called the score a physical fitness score.

At about the same time the Harvard Step Test was further modified for use with *college women*.[42] The bench height was reduced to 18 inches and the exercise duration to four minutes. The scoring formula is the same as for college men.

A further modification in the basic step test was made a number of years later by Brouha and Ball for use with *elementary school boys and girls* below twelve years of age.[43] The bench is lowered to 14 inches. Children under seven years of age exercise only two minutes; children eight to twelve years of age exercise for three minutes. Scoring and classification is the same as for college men in the original test.

Interpretation: Although the step test was originally validated by comparing heart performance with muscular endurance, other investigations have demonstrated that the correlation is low and that heart-lung efficiency seems to be a separate item in over-all physical fitness. Bookwalter found this to be true after comparing Harvard Step Test scores with Army Physical Fitness Test scores.[44] Cureton compared two dozen tests of strength and muscular endurance with Harvard Step Test scores and obtained coefficients of correlation which varied between .002 and .31.[45]

The short form of the test has met with some approval. Karpovich, Starr, and Weiss used it with hospital patients who were permitted to take

[41] Lucien Brouha and J. R. Gallagher, "A Functional Fitness Test for High School Girls," *Journal of Health, Physical Education and Recreation,* 14:517, December, 1943.

[42] Harriet L. Clarke, "A Functional Fitness Test for College Women," *Journal of Health, Physical Education and Recreation,* 14:358–359, September, 1943.

[43] Lucien Brouha and M. V. Ball, *Canadian Red Cross Society's School Meal Study,* Toronto, University of Toronto Press, 1952, pp. 55–56.

[44] Karl W. Bookwalter, *op. cit.*, p. 8.

[45] Thomas K. Cureton, *Physical Fitness Appraisal and Guidance,* p. 288.

mild exercise.[46] Also it appears adaptable for use with college women and high school girls, secondary school boys, and elementary school boys and girls. In the case of elementary pupils this is significant, for few tests of fitness are available for this age group.

TUTTLE PULSE-RATIO TEST

The pulse ratio developed by Tuttle was designed to consider the fine relationship of the resting pulse to the rate following a period of exercise.[47] In this respect it does not differ in principle from the Harvard Step Test and its several modified forms. Over the years the test has been subjected to much research.

Test Equipment Needed: A watch with a second hand, a bench 13 inches high, and a chair to sit on are needed.

Testing Procedure: 1. Record the resting pulse for one minute while the subject is in a sitting position.

2. Instruct the subject to step up and down on the bench for one minute, at a rate of twenty steps per minute for boys and fifteen steps per minute for girls. This is done rhythmically in four counts, with the knees coming to a straight position on the bench.

3. Seat the subject after exercise and immediately count the pulse rate for two minutes.

4. Keep the subject quietly seated until the pulse is back to the resting stage.

5. Repeat the one-minute step-up exercise on the bench. This time the step-up rate is increased to forty for boys and thirty-five for girls.

6. Seat the subject after exercise and immediately count the pulse rate for two minutes.

Scoring: The object is to ascertain the amount of exercise necessary to raise the pulse 2½ times above its resting state; that is, to secure a 2.5 pulse ratio. The following formula is used in computing the pulse ratio:

$$s_0 = \frac{s_1 + (s_2 - s_1)(2.5 - r_1)}{r_2 - r_1}$$

where s_0 = number of steps needed for 2.5 ratio
s_1 = number of steps in first exercise
s_2 = number of steps in second exercise
r_1 = pulse ratio for s_1 $\left(\frac{\text{two-minute pulse count after first exercise}}{\text{resting pulse}}\right)$
r_2 = pulse ratio for s_2 $\left(\frac{\text{two-minute pulse count after second exercise}}{\text{resting pulse}}\right)$

[46] Peter V. Karpovich, Merritt P. Starr, and Raymond A. Weiss, "Physical Fitness Test for Convalescents," *Journal of the American Medical Association*, 126:873, Dec. 2, 1944.

[47] W. W. Tuttle, "The Use of the Pulse-ratio Test for Rating Physical Efficiency," *The Research Quarterly*, 2:5–8, May, 1931.

Example

A subject has a resting pulse of 70, which increases considerably right after the exercise to a total count of 168 in two minutes. Thus, $^{168}/_{70} = 2.40$ (the r_1 ratio). If, after the second bout of exercise, the total two-minute pulse count is 200, then the ratio (r_2) is $^{200}/_{70} = 2.86$. With twenty steps for s_1 and forty steps for s_2, substitution in the formula yields the following:

$$s_0 = 20 + \frac{(40 - 20)(2.5 - 2.40)}{2.86 - 2.57} = 27.2 \text{ steps}$$

The Tuttle pulse-ratio norms indicate thirty-three steps for boys ten to twelve years of age; thirty steps for boys thirteen to eighteen; twenty-nine steps for adult men; and twenty-five steps for adult females An efficiency rating is used to put the scores on a percentage basis. It is based on fifty steps per minute, the performance of well-conditioned athletes in Tuttle's laboratory.

$$\text{Efficiency rating} = \frac{100 \times \text{number of steps obtained for 2.5 pulse ratio}}{50}$$

In the above example an s_0 of 27.2 steps is 54.4 per cent of top efficiency.

To simplify the whole process of testing and scoring Tuttle and Dickinson shortened the test to a single bout of exercise of thirty to forty steps a minute.[48] The shortened version correlated very well (.93 .95) with the original test.

A great many studies, too numerous to mention here, have been carried out involving the Tuttle Pulse-ratio Test and its relationship to physical efficiency, athletic performance, muscular endurance, college women, gymnastics, rope jumping, and the menstrual cycle[49] By and large the findings are favorable to the test This is illustrated in part by the work of Flannigan.[50] It is especially well supported by research workers who have worked closely with physicians and patients with cardiac insufficiency.[51,52,53] The test is a suitable screening instrument to employ to supplement the thorough physical examination

[48] W. W. Tuttle and R. E Dickinson, "Simplification of the Pulse-ratio Technique for Rating Physical Efficiency and Present Condition," *The Research Quarterly,* 11:73-81, May, 1938.

[49] For a more detailed review of the specific research, see the work of John Bovard, Frederick W. Cozens, and E. Patricia Hagman, *Tests and Measurements in Physical Education,* Philadelphia, W. B. Saunders Company, 1949, pp. 73–76.

[50] Kenneth Flannigan, "The Pulse-ratio Test as a Measure of Athletic Endurance in Sprint Running," *The Research Quarterly,* 6:50-52, October, 1935.

[51] Henry Sievers, "A Simple Method of Detecting Abnormal Hearts by the Use of the Pulse-ratio Test," *The Research Quarterly,* 6:36–42, May, 1935.

[52] Ethard N. Lee. "A Further Study of Tuttle's Test as a Means of Detecting Non-compensated Organic Heart Lesions, *The Research Quarterly,* 8:123–128, March, 1937.

[53] Aileen Carpenter, "Further Observations on Tuttle's Test for Non-compensated Heart Lesions," *The Research Quarterly,* 8:130–136, March, 1937.

Suitability of Cardiovascular Measures

Cardiovascular test results are useful as screening measures of physiological fitness. Like tests of muscular strength and muscular endurance they rise with improved personal health status and fall as organic strains and defects appear. These tests, however, require the utmost care in administration in order to secure accurate results.

Usually the subject is required to take a rest period before beginning the test. Repeated tests under similar conditions are sometimes difficult to arrange in the school environment. Furthermore, when the pulse rate is involved, numerous variables such as time of day, body position, and food intake tend to invalidate the measure.[54] Studies indicate that errors in taking the normal pulse count can be great.[55]

Although Gallagher and Brouha and a few others worked with secondary school boys and girls, very little practical application of this type of fitness test has been made in the schools of the country. In fact, there is considerable room for experimentation in the use of tests of cardiovascular endurance in elementary, secondary, and higher education, especially in remedial physical education classes. Certainly pupils with low cardiorespiratory efficiency could profit from this type of screening. This is especially true in those schools where other physical fitness screening tests are not given. Whether a step test or pulse-ratio test is used, pupils scoring poorly could be referred to medical personnel for a more thorough examination.

Probably the chief value of cardiovascular measures is their use today with hospital convalescents and others of low physical fitness. The tests may lose some of their value when they are employed to measure slight variations from normal health status. As physical condition improves to a state approaching that found in the athlete or woodsman, it becomes increasingly difficult for these tests to differentiate between levels of condition.

Apparently, training does not significantly alter the pulse wave sphygmogram of athletes. Because of this Salit and Tuttle concluded that blood pressure measures are of questionable value with healthy young adults.[56] Henry feels that the decrease in heart rate is a more effective test of changes in athletic condition and that the resting heart rate has significant validity as a test of the effect of athletic training.[57] Furthermore, there is a possibility that a two-item test combining heart rate and systolic pulse-wave amplitude

[54] Peter V. Karpovich, *Physiology of Muscular Activity*, Philadelphia, W. B. Saunders Company, 1953.

[55] A. T. Slater-Hammel and L. K. Butler, "Accuracy in Securing Rates by Palpitation," *The Research Quarterly*, 11;90–96, May, 1940.

[56] Elizabeth P. Salit and W. W. Tuttle, "The Validity of Heart Rate and Blood Pressure Determinations as Measures of Physical Fitness," *The Research Quarterly* 15:252–257, October, 1944.

[57] Franklin M. Henry, "Influence of Athletic Training on the Resting Cardiovascular System," *The Research Quarterly*, 25:28–41, March, 1954.

would be a more practical measure of cardiovascular efficiency.[58] Some support is given to Henry's view by Michael and Gallon in a study of college basketball players.[59] A small sample of players was given a one-minute step test at thirty-six steps per minute every three weeks during sixteen weeks of training. The results indicated that the recovery pulse rate made significant changes in three weeks and highly significant changes in six weeks. After six weeks of training, however, a plateau was reached, and no significant changes occurred when training was continued another ten weeks. This appears to illustrate once more the difficulty of differentiating between levels of fitness in well-trained subjects.

Among the numerous tests, the Schneider test, Harvard Step Test and Tuttle Pulse-ratio Test are the most often used. Cureton's research on a number of cardiovascular measures supports the Schneider test, especially if it is given to the subject several hours after eating, when he is at ease.[60] Henry and Herbig agree and make a plea for additional experimental work.[61]

After studying cardiovascular tests in relation to physical performance, it is reasonable to point out again that one reason why heart-lung test scores do not correlate highly with measures of muscular strength and endurance is that they are certainly not the same. They are related but different physical fitness elements. Therefore, a true picture of physical fitness will reflect both cardiovascular and muscular endurance scores.

Beyond the scope of this chapter are more sensitive and more complicated cardiovascular measures, which are used in hospital laboratories and cardiac research centers. They are not suited for school use.

SELECTED REFERENCES

Henry, Franklin M., "Influence of Athletic Training on the Resting Cardiovascular System," *The Research Quarterly,* 25:28-41, March, 1954.

Karpovich, Peter V., Merritt P. Starr, and Raymond A. Weiss: "Physical Fitness Test for Convalescents," *Journal of the American Medical Association,* 126:873, Dec. 2, 1944.

Larson, Leonard A.: "A Note on Scaling Some Measures of Circulation and Respiration," *The Research Quarterly,* 19:290-295, December, 1948.

Massey, B. H., B. F. Husman, and C. L. Kehoe: "The Effect of Posture on the Brachial Sphygmogram as an Indicator of Cardiovascular Condition," *The Research Quarterly,* 28:184-204, May, 1953.

Massey, Benjamin, and Burris F. Husman: "An Exploratory Study of the Ballisto-

[58] *Ibid.*, p. 40.

[59] Ernest D. Michael and Arthur Gallon, "Periodic Changes in the Circulation during Athletic Training as Reflected by a Step Test," *The Research Quarterly,* 30:303–311, October, 1959.

[60] Thomas K. Cureton, *Physical Fitness Appraisal and Guidance,* p. 281.

[61] Franklin Henry and W. Herbig, *op. cit.*, p. 185.

cardiogram as a Measure of Circulatory Fitness," *The Research Quarterly,* 27:316–320, October, 1956.

McCloy, Charles H.: "What Is Physical Fitness," *Journal of Health, Physical Education and Recreation,* 27:201–204, September, 1956.

Michael, Ernest D., and Arthur J. Gallon: "Pulse Wave and Blood Pressure Changes Occurring During a Physical Training Program," *The Research Quarterly,* 31:43–59, March, 1960.

Miller, Waldo H., and Edwin R. Elbel: "The Effect upon Pulse Rate of Various Cadence in the Step Up Test," *The Research Quarterly,* 17:263–269, December, 1946.

Morehouse, L. E., and W. W. Tuttle: "A Study of the Post-exercise Heart Rate," *The Research Quarterly,* 13:3–9, March, 1942.

Phillips, Marjorie, Eloise Ridder, and Helen Yackel: "Further Data on the Pulse-ratio Test," *The Research Quarterly,* 14:425–429, December, 1943.

Pierson, William R., and Philip J. Rasch: "Bruce Physical Fitness Index as a Predictor of Performance in Trained Distance Running," *The Research Quarterly,* 31:77–81, March, 1960.

CHAPTER 7

Measures of Physical Fitness
Strength and Endurance Tests

This chapter discusses measures of muscular strength and muscular endurance with special emphasis on the relationship of these two elements to general health status.[1]

Since human strength and endurance are the foundation of the physical capacity for *all* activity, it follows that no study of physical fitness would be complete without an analysis of these elements and how they may be measured and related to personal health improvement.

Strength Changes and Health Status

From an early age most persons are aware that as their health of the moment deteriorates their physical ability to accomplish simple tasks is noticeably reduced. With the advent of the common cold, overfatigue, inadequate diet, physical strain, diseased tissues, and even a state of mental depression, the functional ability of the muscular system is lowered. Personal efficiency and productivity are thereby decreased. The reverse is also true. When the particular body strain or drain is overcome, capacity for physical activity is often noticeably improved.

The relationship between muscular strength, personal efficiency, and health status is not new. As far back as Brigham, in 1872, physicians measured strength with calibrated dynamometers and expressed the results specifically in pounds or kilograms.[2] Medical doctors, vitally interested in anthropometric measurements, appraised health status by measuring height, weight, girth, breadth, and length of body as well as the strength of the back, legs, chest, and forearm. Dr. Hitchcock, of Amherst College, Dr. Hitchcock, Jr., of Cornell University, and Dr. Sargent, of Harvard University, were among the most prolific experimenters and writers before the turn

[1] See Chap. 6 for a discussion of physical fitness. Included are definitions of physical fitness, motor fitness, cardiovascular fitness, muscular strength, and muscular endurance.

[2] W. W. Tuttle, C. D. Janney, and C. W. Thompson, "Relation of Maximum Grip Strength to Grip Strength Endurance," *Journal of Applied Physiology*, 2:663–670, June, 1950.

of the century.[3] Other leaders such as Hartwell, Anderson, Gulick, Savage, Seaver, and Kellogg did much to show the relationship between strength changes, athletic ability, and all-round personal performance In fact, it was Kellogg who developed a dynamometer for therapeutic exercise, which he called the universal dynamometer because it could measure twenty-five muscle groups.[4]

Although they had wide use, Sargent's original battery of strength tests were poorly organized, were low in reliability, and were difficult to interpret. In 1911, Rogers completed Sargent's Harvard Strength Test as a student at Phillips Exeter Academy Fourteen years later he completed a dissertation at Columbia University in which he improved Sargent's test by standardizing the testing procedure and creating norm tables for age, sex, height, and weight.[5] Rogers' strength index and physical fitness index battery did much to associate strength status with health status Rogers' tests have been used widely in public schools and colleges for several decades, accumulating substantial evidence to support his thesis [6]

> The intimate relations which exist between voluntary muscle activity and the condition and improvement of vital functions, as well as the fact that the chief method of improving physical condition in school is through muscular activity, are the defense for assuming that a measure of muscular efficiency is also a measure of physical fitness.

Rogers, working through the New York State Education Department and later through Boston University, carried on his research by having his graduate students work directly with physical directors and school physicians in the school systems of New York State and New England states By 1940 the research supporting Rogers' thesis was great indeed [7] Since then there has been fairly wide acceptance of Rogers' maxim that "*every change in physical structure, physiological function or psychic state or function is reflected in changed effective-muscle-strength which can be measured*"[8]

[3] See the article by Dudley Allen Sargent, "The Physical Test of a Man" and by Edward Hitchcock, Jr, "A Synoptic Exhibit of 15,000 Physical Examinations," in *Proceedings of the American Association for the Advancement of Physical Education,* Fifth Annual Meeting, Boston, Mass, Apr 4 and 5, 1890

[4] J. H Kellogg, *The Value of Strength Tests in the Prescription of Exercise,* Modern Medicine Library, vol 2, Battle Creek, Mich , 1896

[5] Frederick Rand Rogers, *Physical Capacity Tests in the Administration of Physical Education,* New York, Bureau of Publications, Teachers College, Columbia University, 1925

[6] *Ibid*, p 69

[7] See especially the following reports Frederick Rand Rogers (Ed.), "Boston University Studies in Physical Education," *The Research Quarterly* (Supplement), vol 6, March, 1935, *Fundamental Administrative Measures in Physical Education,* Newton, Mass The Pleiades Company, 1932, Frederick Rand Rogers (Ed), *Education,* vol. 60, April, 1940, Frederick Rand Rogers (Ed), *Education,* vol 60, December, 1940, Frederick Rand Rogers, "The Significance of Strength Tests in Revealing Physical Condition," *The Research Quarterly,* 5 43-46, October, 1934.

[8] Frederick Rand Rogers, "The Significance of Strength Tests in Revealing Physical Condition," *loc. cit.,* p. 46.

Cureton, in 1941, brought together substantial evidence to support strength testing as a way to study physical fitness.[9] Over one hundred references to the literature were listed. During approximately the same period McCloy wrote that " . . . the strength tests in the form of the physical fitness index tell much about the individual's general fitness for living and working."[10] This point was well supported by research at the State University of Iowa.[11]

One of the finest studies to support strength tests as screening measures of physical fitness was carried out at Cornell University, by Chamberlain and Smiley.[12] Here the health service physicians determined health status. There was 80 per cent agreement between their judgment and the physical fitness index scores.

In a long-term study of growth in adolescence, Jones measured four aspects of static dynamometric strength, recorded semiannually for a sample of public school children, aged 11 to 17½ years. Among his many findings was a strong tendency for strength to be associated with a good physique, good health record, early maturing, social prestige, social stimulus value, and an apparently satisfactory level of personal adjustment.[13]

For many years Harrison Clarke has been the champion of strength measures as a means of appraising individual pupil status and giving support to a scientific program of health and physical education. His writings have emphasized the practical application of muscular tests. Clarke's personal ability, initiative, and professional zeal have made strength testing a valuable part of physical fitness measurement. His research at Syracuse University, Springfield College, and the University of Oregon has involved experimentation with subjects in elementary and secondary schools, colleges, hospitals, and rehabilitation centers.[14] In addition to strength testing with dynamometers, Clarke pioneered measuring strength by cable tension. The tensiometer, a refined aircraft control-cable instrument, is used to appraise the strength of individual muscle groups.[15]

[9] Thomas K Cureton (Ed), "Springfield College Studies in Health and Physical Education," *The Research Quarterly* (Supplement), vol. 12, May, 1941.

[10] Charles H McCloy, "How about Some Muscle," *Journal of Health and Physical Education*, 7:302-303, May, 1936

[11] Charles H. McCloy, "Tests of Strength as Measurement of Physical Status," in *Appraising Physical Status*, Iowa City, Iowa, State University of Iowa Studies, vol. 15, no. 2, June, 1938; chap. 7, pp. 60-69

[12] C. G. Chamberlain and D. F. Smiley, "Functional Health and the Physical Fitness Index," *The Research Quarterly*, 2:193–196, March, 1931.

[13] Harold E. Jones, *Motor Performance and Growth*, Berkeley, Calif., University of California Press, 1949. pp. 164 176.

[14] See especially H. Harrison Clarke, "Muscular Strength-Endurance Relationships," *Archives of Physical Medicine and Rehabilitation*, 38:584–595, September, 1957; "Physical Fitness Benefits. A Summary of Research," *Education*, 78:460 466, April, 1958; *Physical Fitness Newsletter*, Eugene, Ore., University of Oregon, Series I to VII, 1953 to 1960.

[15] H. Harrison Clarke, "Comparison of Instruments for Recording Muscle Strength," *The Research Quarterly*, 25:398 -405, December, 1954.

Another form of strength testing was instituted, early in 1950, by Kraus and his associates at the Institute of Physical Medicine and Rehabilitation, New York University. Kraus related substrength behavior to physical inactivity and disease (hypokinetics).[16] He demonstrated that a lack of muscle strength and flexibility accompanies emotional difficulties and the tension syndrome. His minimum strength tests measure the level of strength and flexibility of certain muscle groups.[17] Large numbers of school children have been screened for minimum fitness by the Kraus-Weber battery, both in Europe and the United States.

Physical exercise with its close relationship to muscular energy has been rightly termed "the third phase of medicine."[18] In this respect, the physical education teacher has a real role to play with substrength boys and girls. Gallagher sums this up: "The evaluation of strength, the determination of the disproportion between strength and probable stress, and the increase of strength through exercise can at times constitute better management than a regime which focuses upon the ailment, emphasizes rest, and ignores the facts regarding strength development in exercise."[19]

Strength Changes and Athletic Performance

The literature supports a correlation between muscular strength and athletic ability. Generally speaking, persons capable of highly skilled motor performance in specific athletic behavior are high in physical strength. This is particularly true on an individual basis, that is, when the strength and endurance factor for a particular person is below normal, he is capable of less in physical performance of skill. Moreover, athletic teams have long been equated on the basis of strength tests with resulting game scores indicating a high level of close competition.[20,21,22] Also, personal playing ability in specific sports has been closely related to individual strength and structure.[23]

Just as strength has a bearing on athletic performance, it also affects work efficiency in general. Several studies have illustrated the need for a high

[16] Hans Kraus, Bonnie Prodden, Sonya Weber, and Kurt Hirschham, *Hypokinetic Disease: Role of Inactivity in Production of Disease*, Institute of Physical Medicine and Rehabilitation, New York University Bellevue Medical Center, New York, 1955.

[17] Hans Kraus and Ruth P. Hirschland, "Minimum Muscular Fitness Tests in School Children," *The Research Quarterly*, 25:177–188, May, 1954.

[18] Howard A. Rusk, "Rehabilitation: The Third Phase of Medicine," *Canadian Medical Journal*, 61:603–609, May, 1949.

[19] J. Roswell Gallagher, "Rest and Restriction: Their Conflict with an Adolescent's Development," *American Journal of Public Health*, 46:1424–1428, November, 1956.

[20] Leonard Clark, "Melrose High School Experiments," *The Research Quarterly* (Supplement), 6:111–115, March, 1935.

[21] Carl E. Willgoose, "The Relationship of Muscular Strength to Motor Coordination in the Adolescent Period," *Journal of Educational Research*, 64:138–142, October, 1950.

[22] K. B. DuBois, "Equating Opponents by Means of the Strength Index," *Education*, 60:520–524, December, 1940.

[23] E. Eugene Hooks, "Prediction of Baseball Ability through an Analysis of Measures of Strength and Structure," *The Research Quarterly*, 30:38–43, March, 1959.

strength index when engaging in carrying military packs under military field conditions.[24] The results indicate that strength tests as measures of physical fitness are effective in appraising physical efficiency for military activity.[25]

Finally, because of the volume of research in support of strength testing, it is probable that many more studies in the future will seriously consider muscular strength. Studies in physical reconditioning, sports skill analysis, maturation levels, and body build provide ground for further exploration. Hunsicker and Greey, in reviewing some ninety pieces of research, have shown that work is well under way in a number of these areas.[26] Moreover, strength studies apply to the capacity or efficiency of girls and women quite as much as they do to boys and men.[27]

Grip Strength Testing

The physiology of human muscle and the degree of power that it generates is so sensitive to incoming body stimuli that practically every influence on the organism can be demonstrated through muscle strength changes. Fatigue, weather, time of day, diet, and attitude of mind all affect performance and can be measured.

Because it is difficult to test a person quickly for total body strength, a number of physicians and educators have simply obtained an estimate of strength by measuring the right and left grip strength with a hand dynamometer. Two practicing physicians, Gamboa and Geiss, used grip tests in medical practice as aids to the discovery, prognosis, and treatment of disease. Their findings showed that recovery from illness is accompanied by a recovery in grip strength, which can be carefully plotted on a hospital bedside chart.[28] Blakeman, Jackson, and Rogers demonstrated that the hand dynamometer reflects bodily conditions with considerable accuracy. By checking the grip strength of subjects every hour throughout the day they were able to establish a typical diurnal curve, which closely paralleled similar curves for mental efficiency and organic efficiency.[29] Although there are variations in effective muscular power during the day, and this varies individually, the average person seems to be most effective at about 11 A.M.

[24] H. Harrison Clarke, Clayton T. Shay, and Donald K. Matthews, "Strength Development from Carrying Various Army Packs on Military Marches of 7.5 Miles," *The Research Quarterly*, 26:253 265, October, 1955.

[25] Donald K. Matthews, Clayton T Shay, and H. Harrison Clarke, "Relationship between Strength Loss in Pack Carrying and Certain Motor Physical Fitness Criteria," *The Research Quarterly*, 26:426 439, December, 1955

[26] Paul Hunsicker and George Greey, "Studies in Human Strength," *The Research Quarterly*, 28:109-122, May, 1957.

[27] See especially two reports by Theresa W. Anderson, "Weighted Strength Tests for Prediction of Athletic Ability in High School Girls," *The Research Quarterly*, 7:136–142, March, 1936, 8:69–73, October, 1937.

[28] Frederick Rand Rogers, "The Evolution of PFI Tests and Programs," *Education*, 60:527-540, April, 1940

[29] Ralph N. Blakeman, Perry S. Jackson, and Frederick Rand Rogers, *Further Adventures with Grip Strength Tests*, Newton, Mass., The Pleiades Company, 1934, p. 17.

and 3 P.M.[30] A later study by Wright disclosed a diurnal pattern that is not modified by repeated testing, by starting tests at different times, or by staying awake all night.[31]

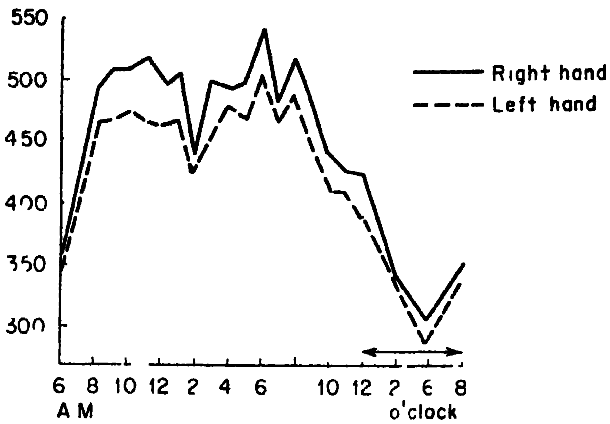


Fig 7-1 The Diurnal Curve for Grip Strength in a Twenty-four-hour Period. SOURCE: Verna Wright, "Factors Influencing Diurnal Variations of Strength of Grip," *The Research Quarterly,* 30.110–116, March, 1959

Grip strength testing should not be considered a substitute for a more thorough measurement of strength. It is simply a screening device that notes the immediate reaction of muscles to some variable or power change for any one person Willgoose used grip strength to measure changes due to tobacco smoking.[32] So did Kay and Karpovich.[33] Slater-Hammel discovered that exercise of one arm in flexion and extension produces a positive and significant improvement in the muscular performance of the other.[34] Tuttle, Janney, and Thompson showed that a person with a stronger grip can maintain a higher level of strength for a period of effort, the result being the same for both hands.[35] Everett and Sills related grip strength scores to somatotype, stature, and various measures of the hand. They found that weight, hand width, height, and mesomorphy (athletic build) were the most influential variables in the prediction of grip strength [36] Cousins found that grip strength tests were reliable and that trained and untrained testers were equally effective in administering tests [37] Because the smallest mean grip

[30] *Ibid*, p 16.

[31] Verna Wright, "Factors Influencing Diurnal Variation of Strength of Grip," *The Research Quarterly,* 30:110-116, March, 1959.

[32] Carl E. Willgoose, "Tobacco Smoking, Strength and Muscular Endurance," *The Research Quarterly,* 18·219-225, October, 1947.

[33] Hector W Kay and Peter V Karpovich, "Effect of Smoking upon Recuperation from Local Muscular Fatigue," *The Research Quarterly,* 20 250–256, October, 1949.

[34] Arthur T. Slater-Hammel, "Bilateral Effects of Muscular Activity," *The Research Quarterly,* 21:203-209, October, 1950

[35] W. W. Tuttle, C. D Janney, and C W. Thompson, *op cit*, p 670.

[36] Peter W. Everett and Frank D. Sills, "The Relationship of Grip Strength to Stature, Somatotype Components, and Anthropometric Measurements of the Hand," *The Research Quarterly,* 23:161–166, May, 1952.

[37] George F. Cousins, "Effect of Trained and Untrained Testers upon Administration of Grip Strength Tests," *The Research Quarterly,* 26:273–278, October, 1955.

strength scores are found in the youngest age group and the greatest scores in the older age group, Bookwalter established grip strength norms for males.[38] These are in the form of achievement scales for each hand and age from nine to twenty-four years and up for weight groups from 69 to 210 pounds and up; and for Classification Index I groups. Bookwalter has attempted to control the factor of individual differences by treating grip scores in terms of sex, age, weight, and Classification Index, thus a person can be compared with his peers. Probably the chief reason there is not a widespread use of these norm tables is that a more thorough measurement of body strength is being employed. Furthermore individual grip strength seems to be such a variable item that a person might best be compared only with himself as succeeding measurements of grip strength are taken.

THE STRENGTH INDEX

One of the chief difficulties in using physical performance to appraise muscular strength changes is the degree of skill that influences the score. Rogers in explaining his strength index points out that:[39]

> the chief virtue of the tests . . . are that the qualities they measure are not greatly susceptible of improvement by increase in skill alone . . . that is, they are as nearly tests of physical capacity exclusively as it is possible to devise. To be substituted for the strength tests described . . . any new battery should be proven more valid, more reliable, more objective, more economical, possess greater total utility, not be susceptible of improvement because of increased skill alone, and be supplied with properly determined norms for all combinations of sex, weight and age.

The strength index (SI) is the gross score obtained from a battery of six strength tests in pounds and a measurement of lung capacity in cubic inches. The six tests are right grip, left grip, back lift, leg lift, pull-ups, and push-ups. The strength index is intended primarily as a measure of general athletic ability. The previous discussion supports this intention. It can be used as a measure of physical fitness for both sexes, ages eight to thirty-eight, especially when employed as an integral part of the physical fitness index (PFI).

Test Equipment Needed. The following items are needed: wall scale or stadiometer for measuring height, weight scales, wet spirometer with wooden mouthpieces for measuring lung capacity, and a manuometer for measuring grip strength[40], score cards and pencils, a back and leg dynamometer, calibrated in pounds and capable of registering a 2,500-pound lift (the chain to

[38] Karl W. Bookwalter, "Grip Strength Norms for Males," *The Research Quarterly*, 21:249-273, October, 1950.

[39] Frederick Rand Rogers, *Fundamental Administrative Measures in Physical Education*, loc. cit., p. 109.

[40] This equipment can be obtained from Naragansett Gymansium Equipment Co., Centralia, Mo. The Fred Medart Products, Inc., 3535 DeKalb St., St. Louis, Mo., supplies back and leg dynamometers with lifting belt.

the dynamometer should be no less than 25 inches in length and the lifting bar to go with it should measure about 20 inches in length); a 5½-foot lifting belt about 4 inches wide and $\frac{5}{16}$ inch thick should be made of tightly woven canvas, pliable and similar to fire hose material (there is a loop at the end of the belt big enough to slip over one end of the lifting bar); block of magnesium carbonate to prevent slipping of hands on the lifting bar; high horizontal bar for pull-ups (boys); low horizontal bar adjusted for height (girls); floor mat to prevent slipping; gymnasium parallel bars or wall parallel bars for push-ups (boys), and gymnasium stall bar bench, 13 inches high by 14 inches wide by 20 inches long, for push-ups (girls).

Testing Procedure: This test can be given by one person to another, or it can be a group test. In the latter case it will be necessary to have a tester at each testing station. For a large group of pupils to be processed quickly and efficiently a total of eight testers at eight testing stations is recommended. A single person can be tested in about ten to twelve minutes, but in a group approximately sixty persons per hour can be measured. The gymnasium is an excellent location to carry out the test. Trained testers are necessary, especially for the administration of the back and leg dynamometer tests.

The order of administering the strength index battery is as follows:

1. Record *name, date,* and *age* in years and months on a physical fitness index (PFI) score card.

DEPARTMENT OF ATHLETICS AND PHYSICAL EDUCATION, SYRACUSE UNIVERSITY

(Form 120-26) SU TM 1-47—10154 (41a) MINIMUM ESSENTIALS TEST RECORDS

Date							SWIMMING	
Age	Y M	Y M	Y M	Y M	Y M	Y M		
Weight								
Height	In.	In.	In.	In.	In.	In.	SKILL TESTS	
Multiplier								
Pull-ups								
Push-ups								
Arm Strength								
Leg Lift								
Back Lift								
Left Grip								
Right Grip								
Lung Capacity								
Strength Index								
Normal S. I.								
P. F. I.								

Name .... College .... Year ....

Fig. 7-2. Score Card for strength index and PFI.

2. Record *height* to the nearest half inch and *weight* to the nearest pound with subject in gym suit with shoes removed.

3. Administer test of *lung capacity,* recording wet spirometer measurement in cubic inches. Proceed as follows:

Stand the wet spirometer on a high table so that the subject blowing into the tube can stand erect when beginning the test. Make certain the spirom-

eter is filled with water within one inch of the top Have the subject insert a clean wooden mouthpiece into the rubber tube (tube 36 to 42 inches long)

a Have the subject take one or two deep breaths before the test Then have him inhale as fully as possible place the mouthpiece in his mouth, and slowly and steadily expell all the air in his lungs He should blow every bit of air out, bending slowly forward as he finishes

b The tester should watch carefully the cubic inch indicator as it rises on the spirometer so that its highest point can be noted

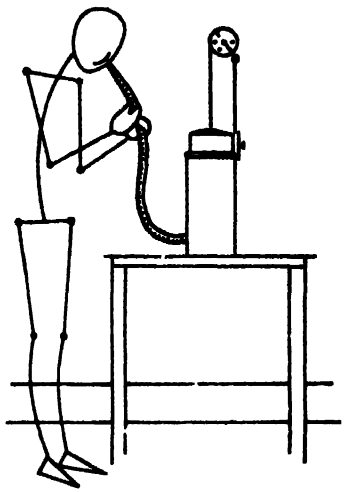


Fig 7 3 Testing for Lung Capacity

c Remove the spirometer base plug and permit the upper container to descend slowly to normal so that the next subject can be tested The test should be repeated for the same subject if the first test did not represent his best effort A best effort will usually be one in which a full inspiration and full expiration is made with no air escaping either through the nose or around the mouthpiece

4 Test *grip strength* (right hand)

A rectangular type hand dynamometer is placed face down in the palm of the subject's hand so the wide edge meets the fingers and the rounded edge meets the base of the hand When the subject squeezes the dynamometer the thumb will usually touch or overlap the first finger [11]

a Place the dynamometer in the subject's right hand Tell him to swing the arm downward and squeeze the hand dynamometer as hard as he can

[11] Evidence is available to show greater test reliability and results when the face of the dynamometer is placed down in the palm of the hand See research by D B Van Dalen and C A Peterson, "A Comparative Study of the Administration of the Manuometer,' *The Physical Educator,* 7 52–58, May 1950 C L Wear, "Further Study of the Administration of the Manuometer," *The Physical Educator,* 9 82–86, October 1952

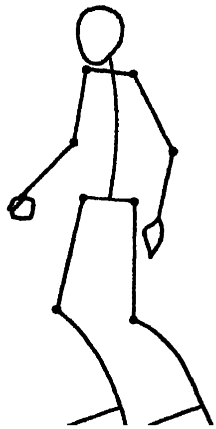


Fig. 7-4 Grip Strength Testing with Manuometer.

Be careful that he does not touch his body or some nearby object. If he does, he should have another test within a half minute.

b. After the test, read the score to the nearest pound and return the indicator to zero.

5. Test *grip strength* (left hand).
6. Administer the *back lift* test

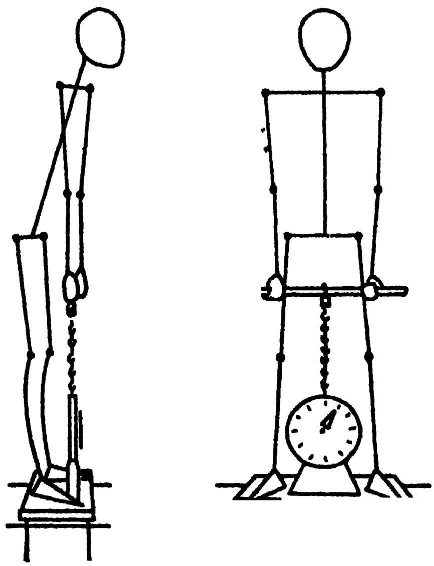


Fig 7-5. The Back Lift Test

The back and leg dynamometer should be calibrated from 0 to 2,500 pounds. It should be anchored to a wooden base big enough to stand on, and it should have some freedom to move forward, back, and laterally. The dynamometer with its base should be placed on a small elevated platform about chair high where the tester can work comfortably in measuring the subject. With a solid base the person being tested will feel secure and

comfortable while lifting. When lifting, the subject should straddle the dynamometer with his feet parallel and about 6 inches apart. The length of the chain from the dynamometer to the lifting bar should be about 25 inches long. The iron handle with its lifting hook in the center should be about 20 inches long. Standardized equipment usually includes all these features. The handle can be taped to permit a more secure grip.

a. Have the subject stand in an erect position on the base of the dynamometer with his fingers extended downward on the front of the thighs.

b. Hook the chain to the lifting bar at a point just below the fingertips.

c. Then have the subject grasp the bar firmly with both hands: one with an overhand grip and the other with an underhand grip, and "lock" his knees in a perfectly straight position. With his eyes looking straight ahead, he is now in the position to perform the back lift.[42]

d. Grasp the subject's hands firmly to hold them to the bar. A little magnesium carbonate chalk rubbed on the hands will remove any moisture that may cause the hands to slip. Instruct the subject to lift, steadily pulling as hard as he can with his back. At the end of the lift the subject's back should be almost straight.

e. Read the indicator on the dynamometer and record score.

7. Administer the *leg lift* test.

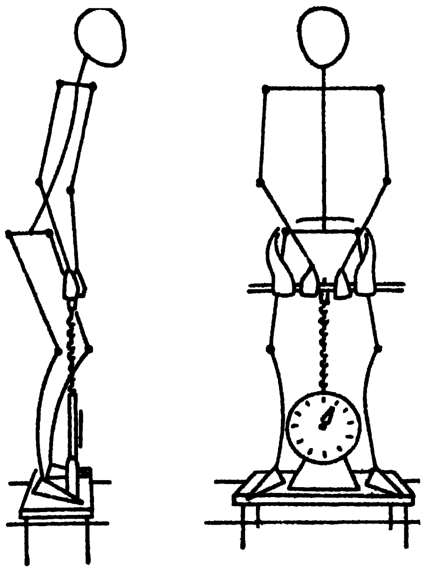


Fig. 7-6. Testing Leg Strength.

Although a number of physical instructors still test leg strength without the use of a belt, the vast majority have adopted the new technique and use

[42] When the grips are reversed and the subject is in the slightly bent position and performs a lift, he will not be able to straighten up completely. Here the leverage is such that maximum power can be applied to the instrument, and it is very difficult to strain the back muscles.

the belt norms (see Tables A-26 to A-29, Appendix A) [43] The belt, a heavy, pliable canvas material should have a small loop on one end This loop should be tightly secured to the canvas with a metal bolt so that it will not be easily disturbed when stress and strain is applied It slips over one end of the iron lifting bar

a Have subject stand on the base of the dynamometer and hold the lifting bar near the center with an overhand grip so the bar rests at the junction of the upper thigh and abdomen

b Face the subject Loop the belt around one end of the bar Then carry the belt behind the subject and across his hips looping the free end around the other end of the bar Tuck the free end under so that it rests against the body This will provide a tight secure arrangement for lifting With practice this operation can be performed very easily

c Take hold of the dynamometer chain and ask the subject to bend his knees (squat position) slightly to an estimated 115 to 124 degrees After several leg tests have been administered it will be relatively easy for the tester to estimate the lifting angle that will permit maximum lifting power Practice will prove helpful

d Make sure that the subject's knees are straight ahead his arms and back straight and his head erect with his eyes looking straight ahead

e Instruct the subject to lift driving his legs downward as hard as he can A maximum lift will occur when the subject's legs are almost straight at the end of the lifting effort If the legs fail to approach an *almost* straight position or if the knees actually 'lock or straighten out a second lift should be taken a half minute later Generally the leg lift will be about 2½ to 3 times as great as the back lift

f Read the indicator on the dynamometer and record score

8 Measure arm strength by *push-ups*

Boys Either gymnasium parallel bars or wall parallel bars can be used to measure the number of times a person can push up (dips

a Have the subject stand at the end of the bars with his hands placed on the bars

b Instruct him to jump to the front support position with arms straight Count this as his first push-up He should then perform as many push-ups as he can, quitting only when he can no longer push In doing this the proper manner is to lower the body until the angle of the upper arm and forearm is a right angle or less From here the subject pushes up to the straight-arm position He should do these push-ups without kicking or swinging

c Count each push-up from the lower position When a subject fails to lower his body to the proper bent-arm angle, credit him with only half a push-up

[43] The belt technique is by far the most valid and objective means of measuring leg strength See Edgar W Everts and Gordon J Hathaway, "The Use of the Belt to Measure Leg Strength Improves the Administration of Physical Tests," *The Research Quarterly,* 9 62–66, October, 1938.

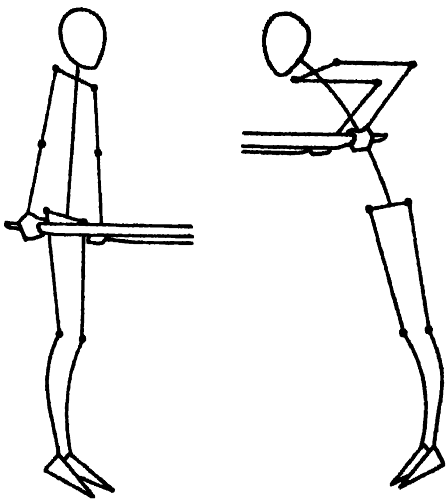


Fig. 7-7. Boys' Push-ups

d. Total the push-ups and record.

Girls: Push-ups are performed on a small, 13-inch-high stall bar bench. The girl grasps the edge of the bench in a front leaning support position. This is the starting position. The feet are together with the balls of the feet on the mat; the body is held straight with head up; the arms are straight and at a right angle to the body.

a. Have the subject lower the body so the chest touches or nearly touches the bench.

b. Each push-up from the bench counts one. A girl is permitted four half counts for incomplete movements or arching and swaying the back.

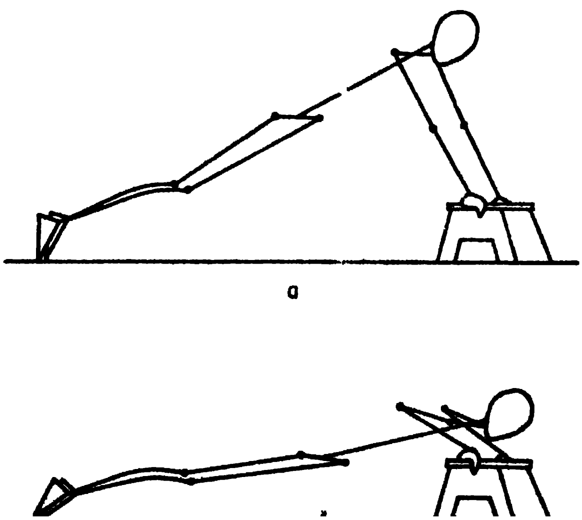


Fig. 7-8. Girls' Push-ups.

10 Measure arm strength by *pull-ups*

Boys: Either rings or a horizontal bar can be used to measure the number of times a subject can chin himself These should be high enough for the tallest person to hang without his feet touching the floor

a Have the subject hang from the rings or bar using an overhand grip When perfectly still, ask him to chin himself as many times as he can, quitting only when he cannot possibly pull up another time On each chin-up the chin should come all the way to the level of the hands Each time he drops down his arms should be perfectly straight before chinning again Do not permit kicking or swinging

b. Count each pull-up from the straight-arm position When a subject fails to lower his body all the way between pulls or only pulls up halfway, count half a pull-up.

c. Total the pull-ups and record

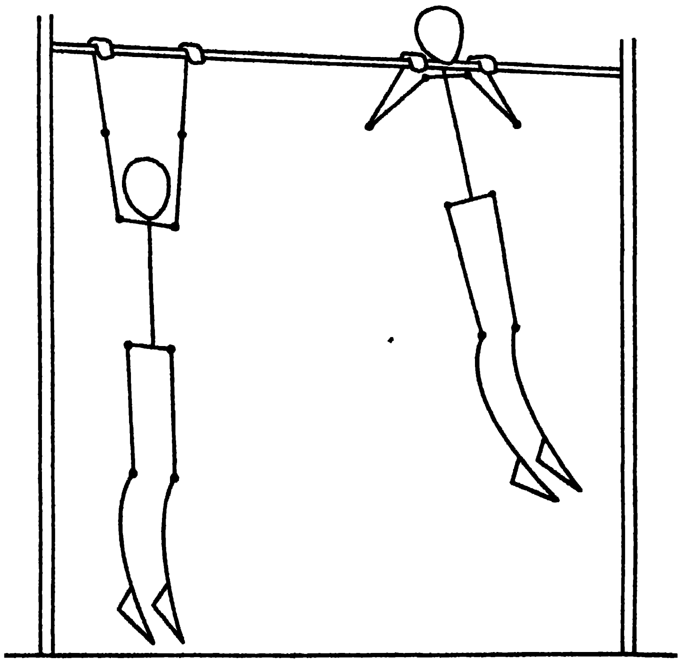


Fig 7-9 Pull-ups for Boys

Girls: Although pull-ups can be done on the low horizontal bar, a pair of adjustable rings permit a greater freedom of action. These rings can be attached to a horizontal bar or to one bar on a set of parallel bars A gymnasium mat is placed on the floor to prevent the feet from slipping [44]

a. Adjust the rings to the subject's apex level of the sternum

b. Have the subject grasp the rings and lower herself under the bar until

[44] A platform with an adjustable heel rest can be substituted for the mat The rest can be raised or lowered according to the height of the girl being tested, thus permitting the rings to remain at a fixed height. This device is distributed by A E. Gay, Public Schools, Lockport, N.Y.

the arms are completely straight forming a right angle with the body. The body is also held straight with the weight resting on the heels.

Testing Suggestions: 1. In order to standardize testing procedures for pull-up and push-up tests the tester should count so that the subject can hear how well he is doing; when half-counts are given, the reason should be quickly explained. Also, there should be a five-minute rest between push-ups and pull ups but no rest between the other test items.

2. As in all physical education activity, a medical examination should precede the test.

3. The persons taking the test not only should know what is being appraised but also should have some familiarity with the test items. Test results are more reliable when a person has had some experience with the test items.

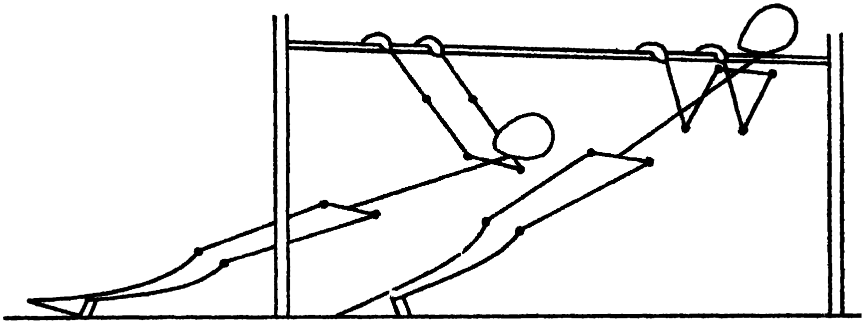


Fig 7-10 Pull-ups for Girls

4 The subject should always be encouraged to give his best effort. This is obviously necessary for the validity of any test, but it is not proper to exhort the subject to greater effort through undue motivation

5. It is advisable to have a little warm-up activity in the form of trunk bending and toe touching for the back lift test and three or four deep knee bends preceding the leg lift

6 Trained testers, especially for the back and leg lifts, are very important for test reliability Also, the accuracy of test instruments should be checked periodically

Scoring: The strength index (SI) in the total test score is calculated by adding together the scores made on each test item: lung capacity, right grip, left grip, back lift, leg lift, and arm strength.

Arm strength is determined by noting the number of push-ups and pull-ups recorded Arm strength is figured according to the formula

$$(\text{Push-ups} + \text{pull-ups}) \times \left(\frac{W}{10} + H - 60\right)$$

where W = weight, pounds

H = height, inches

Fractions are corrected to whole numbers.

Example

A boy who weighs 140 pounds and is 66 inches tall does 8 push-ups and 10 pull-ups. According to the formula, $(8 + 10) \times (^{140}/_{10} + 66 - 60) = 18 \times 20$, which gives an arm strength of 360 pounds.

Interpretation The strength index is a major part of the physical fitness index. By itself it is a measure of power and correlates well with general athletic ability. By and large, boys and girls who have the strength, endurance, speed, and agility to perform well in the various physical activities associated with sports and athletic skills also have high SI scores.

There is a high enough correlation between SI test items to indicate a satisfactory contribution of individual test items to the test battery as a whole. This has been demonstrated by several investigations using these tests. Sambolin for instance found intercorrelations from .06 (arm strength and lung capacity) to .67 (leg strength and back strength).[15] Had the correlation been higher the argument might be advanced that one test should be used as a substitute for the total battery.

It has been suggested from time to time that the measurement of vital capacity has little to do with bodily strength and should be eliminated from the strength index battery. It is true that lung capacity is not a test of strength. But it is further true that changes in lung capacity occur with changes in strength and health status.[16] Dail and Affeldt found that vital capacity is related to strength endurance, heart and lung disease, and general body weakness.[17] Clarke related lung capacity to maturity, age, height, weight, cable-tension strength and McCloy's Classification and Athletic Indices and obtained correlation coefficients varying from .80 to .86.[18] Also, correlations as high as .61 between lung capacity and swimming time in the 200-yard swim have been noted by Davis.[19] Over the years users of the strength index and physical fitness index have found that a drop in lung capacity generally accompanies a drop in muscular strength. Because of this, lung capacity scores afford one more means of screening pupils for poor health. These points together with the fact that the Physical Fitness Index Test norms were established with the lung capacity test included, afford a partial argument for leaving this measure in the strength index battery.

[15] Luis Sambolin, "Extent of Relationship between Several Selected Strength and Cardiovascular Tests," unpublished master's thesis, Syracuse University, Syracuse, N.Y., 1943.

[16] Samuel A. Weisman, *Your Chest Should Be Flat*, Philadelphia: J. B. Lippincott Company, 1938, pp. 30-42.

[17] Clarence B. Dail and John E. Affeldt, "Vital Capacity as an Index of Respiratory Muscle Function," *Archives of Physical Medicine and Rehabilitation*, 37:383-392, June 1957.

[18] H. Harrison Clarke, "Relation of Physical Structure to Motor Performance of Males," *Contributions*, American Academy of Physical Education, 1958, no. 6.

[19] Jack F. Davis, "Effects of Training and Conditioning for Middle Distance Swimming," unpublished doctoral dissertation, University of Oregon, Eugene, Ore., 1955.

THE PHYSICAL FITNESS INDEX

The physical fitness index (PFI) is a measure of physical condition. It is calculated by dividing achieved strength index by a *norm* for sex, age, and weight at the time of testing. Norms are available for ages eight to thirty-eight for boys, girls, men and women. These can be obtained from the Syracuse University Press, Syracuse, New York. See Tables A-26 to A-29, Appendix A, in order to determine the normal strength index for a particular person.

DEPARTMENT OF ATHLETICS AND PHYSICAL EDUCATION SYRACUSE UNIVERSITY
MINIMUM ESSENTIALS TEST RECORDS

Date	10/14/61	4/0/62					SWIMMING
Age	Y 18 M 1	Y 18 M 7	Y M	Y M	Y M	Y M	
Weight	118	122					
Height	64 in	64 in	in	in	in	in	SKILL TESTS
Multiplier	16	16					
Pull ups	12	14					
Push-ups	8	10					
Arm Strength	320	384					
Leg Lift	680	900					
Back Lift	320	350					
Left Grip	68	72					
Right Grip	74	78					
Lung Capacity	192	220					
Strength Index	1654	2004					
Normal S I	1877	2038					
P F I	88	98					

Name THOMAS T JONES College EDUCATION Year 1964

Fig. 7-11 PFI Test Scores

In Figure 7-11 the boy had a SI of 1654 in October. In comparison with the norm for his weight, sex and age he is substrength. His PFI is 88. This was obtained by the following formula:

$$\text{PFI} = \frac{\text{achieved SI}}{\text{normal SI}} \times 100 = \frac{1654}{1877} \times 100 = 88$$

Another PFI test for this boy six months later showed some improvement. His PFI increased to 98.

Interpretation and Use of PFI Tests

The strength index is limited chiefly to comparing subjects with themselves, but the PFI can comparing subjects with each other. This is possible because the strength factor of any one person is related to a norm table. Without norm tables two subjects of equal SI might appear the same in potential output. In fact, they might appear quite alike in their ability to exert great power for the 100-yard dash, even though their age and weight might differ. However, it is highly unlikely that they would be the same in

endurance, a quality calling for sustained effort. Weight, in this case, is a limiting factor. Therefore, a boy with an SI of 1800, weighing 150 pounds, would be quite unlike another boy with an SI of 1800, weighing 120 pounds. Other things being equal, the lighter boy would have the higher PFI score and be capable of prolonged effort and greater physical and mental endurance. He would recuperate faster from fatigue and do more work in a given time than a person whose PFI was lower.

There are some who argue that by giving PFI tests and emphasizing evaluation of physical fitness there is a tendency to slight other objectives of the physical education program. Nothing could be further from the truth. One has only to study the many school systems from New England to the West coast to see the close relationship between physical fitness activities and such items as sports skills, the competency of athletic teams, the extent of the intramural program, and the development of social efficiency and recreational competency. No program in which physical fitness is properly appraised and developed has slighted other goals of physical education. The very opposite is the rule.

The Physical Fitness Index Test has been used extensively for individual health guidance. Each subject with a PFI below 85 is considered separately. Often the case-study approach is used to determine the underlying cause of the low score. See Chapter 16 for a broad discussion of how test scores can be used to meet individual needs.[50] The lower the PFI, the greater the need for rest, special exercise, diet, and advice by physicians and other health personnel. Low PFIs are the results of physical defects, improper food or eating habits, improper exercise, emotional disturbances, or other physical or mental strains, drains, or injuries. Persons with PFI below 85 need guidance to engage profitably in supervised remedial activities or rest daily, until their index is 100 or higher. Otherwise they are likely to decline steadily to even lower levels of physical fitness.

Although the low PFI is the chief concern of health personnel, the subject with the very high PFI also bears watching Since persons with PFI in the 100s have greater physical and mental endurance, they frequently drive themselves into states of chronic fatigue, sleeplessness, and irritability. They become examples of the tension syndrome so common in the hustle and bustle of Western civilization. Zimmerli showed that a number of subjects with PFI scores from 156 to 202 exhibited characteristics of the high-strung, overambitious person who was unable to relax and take life in stride.[51]

It cannot be emphasized too strongly that *repeated PFI tests* determine whether a subject is gaining or losing in physical fitness and therefore whether a remedial regime is adapted to the subject's needs. Obscure de-

[50] See also Carl E. Willgoose, "Meeting Individual Physical Fitness Needs," *School and Society*, 67:444–447, June, 1948.

[51] Elizabeth Zimmerli, "Case Studies of Unusual Physical Fitness Indices," *The Research Quarterly* (Supplement), 6:246–248, March, 1935.

fects and wrong personal habits, which may escape the attention of health service personnel, often are discovered through comparison of PFI records. It is readily possible for persons of almost any age to raise their PFI score from 40 to 100 per cent, with proportionate increases in personal fitness and happiness.

As PFI scores improve, especially at the lower levels, so does mental capacity for the intellectual activity. Simply stated, one is more *able* to reach his innate potential as expressed by IQ and other aptitude-for-learning scores. Rogers has demonstrated this relationship by comparing development of physical fitness with improvement in scholarship. In fact, Rogers contends that general learning potential is about twice as dependent on health status as it is on intelligence; thus he postulates:

$$\text{General learning potential} = (\text{IQ})(\text{PFI})^2$$

Certainly, sick people cannot think up to par, or above par. However, what remains to be shown experimentally is the *degree of fitness* necessary to improve mental output at a given intelligence level. This is a rich area for new research.

Modifications in PFI Tests

Over the years a number of changes in norms, administration, and use have been made for the PFI battery These have improved it measurably as a screening device in the school health program.

McCloy and Young advocate several revisions in the strength index.[52] One is changing the arm strength formula from $(W/10 + H - 60)$ to

$$\text{Boy's arm strength} = 3.54\ (\text{weight}) + 3.42\ (\text{push-up} + \text{pull-up}) - 92$$

For example, by applying th. change to the second PFI test recorded in Figure 7-11, an arm strength score of 422 would be derived:

$$3.54\ (122\ \text{pounds}) + 3\ 42(10 + 14) - 92 = 422$$

This figure, when compared with the Rogers arm strength formula of 384, is definitely higher and would give the subject a higher SI and slightly higher PFI. McCloy reasons that Rogers' formula penalizes the small boy, so he emphasizes individual body weight McCloy also has a formula for *either* push-ups or pull-ups for boys, which is calculated as follows:

$$1.77\ (\text{weight}) + 3\ 42\ (\text{push-ups } \textit{or} \text{ pull-ups}) - 46$$

The formula for girls is slightly different:

$$\text{Push-up strength} = .78\ (\text{weight}) + 1.1\ (\text{push-ups}) + 74$$
$$\text{Pull-up strength} = .67\ (\text{weight}) + 1.2\ (\text{pull-ups}) + 52$$

[52] Charles H. McCloy and Norma D. Young, *Tests and Measurements in Health and Physical Education,* New York, Appleton-Century-Crofts, Inc., 1954, p. 129.

In an effort to determine the validity of McCloy's formula Clarke correlated the body weights of secondary school boys and college men. The obtained correlation of .95 was due to the extra weighting in the McCloy formula.[53] Matthews and Golnick went further in seeking validity by comparing the two arm strength formulas with energy cost as the criterion measure.[54] Their findings revealed that the energy cost of the activity was not significantly related to either biceps girth or total number of push-ups plus pull-ups. Although McCloy's formula had a correlation of .40 with energy cost and .33 with weight, neither formula could be employed to any significant extent as a predictive measure.

McCloy also worked out an *athletic strength index* and a *pure strength index* by putting the SI test items into the following formulas[55]:

Athletic strength index:

Right grip + left grip + .1 (back lift) + .1 (leg lift) + 2 (pull-up strength) + push-up strength − 3 (weight)

Alternate form: same as above minus back and leg lifts. This measure correlates .91 with a criterion of six track and field events. Norms are available in McCloy text.

Pure strength index:

5 (right + left grip) + .1 (leg lift) + pull-up strength + push-up strength

"Pure" strength and body size are factors considered by this index. Norms are not available.

The Oregon Simplification. Clarke and Carter experimented with chronological age and eight anthropometric tests in an effort to simplify the SI and PFI.[56] They discovered that weight and age had the highest multiple correlation with SI; that Rogers norms are still adequate (having been compared experimentally with test scores for secondary school boys in eleven widely dispersed areas of Oregon): that lung capacity, pull-ups, and leg lift of elementary school boys correlated .87 with PFI; that push-ups, pull-ups, leg lift, and lung capacity correlated .85 with PFI in junior high school boys; and that the simplification can be given more rapidly and with fewer testers than the complete test. Regression equations were derived, and formulas for the shortened PFI test were set up as follows:

Upper elementary school boys:

SI = 1.05 (leg lift) + 1.35 (back lift) + 10.92 (push-ups) + 133

[53] H. Harrison Clarke, "Relationship of Strength and Anthropometric Measures to Various Arm Strength Criteria," *The Research Quarterly,* 25:134–143, May, 1954.

[54] Donald K. Matthews and Philip Golnick, "Energy Cost of Pull-ups and Push-ups as Related to Arm Strength Formulas," *The Research Quarterly,* 30:292–296, October, 1959.

[55] Charles H. McCloy and Norma D. Young, *op. cit.,* pp. 25–26.

[56] H. Harrison Clarke and Gavin H. Carter, "Oregon Simplification of the Strength and Physical Fitness Indices," *The Research Quarterly,* 30:3–10, March, 1959.

Junior high school boys (alternate formulas):

A. SI = 1.33 (leg lift) + 1.20 (arm strength) + 286
B. SI = 1.12 (leg lift) + .99 (arm strength) + 5.19 (right grip) + 129

Senior high school boys (alternate formulas):

A. SI = 1.22 (leg lift) + 1.23 (arm strength) + 499
B. SI = 1.07 (leg lift) + 1.06 (arm strength) + 1.42 (back lift) + 194

Here, arm strength is figured by the Rogers method. Also, the A test is the shorter one, with the B test, therefore, coming closer to the boy's actual SI. In order to compute the PFI under the Oregon simplification method it is necessary only to substitute actual test scores in the formula and consult the Roger's norm table for the normal SI.

Example

Consider an elementary school boy:

Leg lift = 700	Age = 11 years 6 months
Back lift = 250	Weight = 110 pounds
Push-ups = 4	

SI = 1.05(700) + 1.35(250) + 10.92(4) + 133 = 735 + 337 + 44 + 133 = 1249

The norm for age and weight is 1242, obtained in the usual manner, so the PFI would be

$$\frac{\text{SI } 1249}{\text{Norm } 1242} \times 100 = 100$$

Adaptation to Norms

Norm tables are never perfect; they exist for purposes of comparison, assuming that persons are about equal in the characteristics being considered. Unfortunately, no two persons are the same, which can only indicate that norm tables must be used in the light of other known differences or conditions. For example, PFI norm tables were constructed using subjects with various extremes of body build (constitutional types). Somatotypically, extreme endomorphs were grouped with extreme ectomorphs and physically active mesomorphs. Thus the resulting norm tables represent a conglomeration of averages of body build. Therefore, the heavy, roly-poly, endomorphic boy, whose structure is normal enough to him, is penalized by having his score related to a norm table made up of the body types more favorable to strength and endurance appraisal. Willgoose and Rogers found that the more endomorphic a person is, the lower his PFI score[67]; also the more

[67] Carl E. Willgoose and Millard L. Rogers, "Relationship of Somatotype to Physical Fitness," *Journal of Educational Research*, 62:704-712, May, 1949.

mesomorphic, the greater the PFI. Moreover, where the predominantly mesomorphic person had an endomorphic component of less than 4 in his somatotype, the average PFI score was 107; however, when the endomorphic component was 4 or more, the mean PFI score dropped to 81. This is further evidence of the effect of endomorphy on bodily strength.

This research suggests that somatotype should be considered when PFI scores are being evaluated. There are a number of PFI score cards with a space provided for somatotype rating. In interpreting the score, a PFI of 80 for an extreme endomorph would probably be far less serious than it would for a mesomorph or ectomorphic combination. The latter types, where weight is not a handicap, are generally expected to put in a performance close to or above the norm.

LARSON MUSCULAR STRENGTH TEST

Using factor analysis, Larson and Yocum studied the muscular strength tests of Rogers and McCurdy and found that a combination of pull-up strength, push-up strength, and the vertical jump gave as valid results as a complete strength test.[58] Dynamic strength and static strength were two of the factors isolated. Since the dynamic strength test items proved three times more valuable in predicting general motor ability than static strength items, Larson built his test for high school and college men to include the following:

1. *Pull-ups* on horizontal bar. Subject starts from a hanging position and continues as long as possible. To score, count one point for each full chin.

2. *Push-ups* on end of parallel bars. Subject has ten-minute rest between pull-ups and push-ups. To score, count one point for each full dip.

3. The *vertical jump* test is given three times to get best score.

a. Subject stands facing wall as close as possible, with both arms extended upward, feet flat on floor. Where the fingers make their maximum reach, a chalk mark is made and a line drawn at this point.

b. Subject stands with one side to the wall. Before subject jumps, chalk dust is placed on the middle finger of hand on wall side.

c. The subject jumps as high as possible and touches wall board with chalked finger. The distance between the reach (chalk line) and jump mark is recorded to the nearest half inch.

4. To score:

a. Change raw scores to weighted standard scores. See Table A-16, Appendix A.

b. Total the weighted standard scores.

[58] Leonard A. Larson and Rachel D. Yocum, *Measurement and Evaluation in Physical, Health, and Recreation Education*, St. Louis, The C. V. Mosby Company, 1951, pp. 86, 97.

c. Apply obtained score to Larson's Classification Index:

Excellent	360–up
Good	309 360
Average	249 308
Poor	190 248
Very poor	189 -down

Example

A high school boy doing 10 pull-ups, 10 push-ups, and completing a 20-inch vertical jump would score 119, 52, and 152, respectively This total of 323 points would be considered a good score

This test has been included here because it is essentially a strength and power test, with an extremely low minimum of skill being employed in the performance It can be used, therefore, as a screening measure for physical fitness status and in work with substrength persons It also can be used in predicting general motor ability

CABLE-TENSION TEST

Orthopedic difficulties have long been appraised by physical medicine personnel using various manual tests of muscle strength and endurance The validity of such testing depends upon a high degree of skill and experience in the person administering the tests Daniels, Williams, and Worthington produced a guide to muscle testing that standardized the testing technique Since no calibrated instruments were used in the procedure, there remained considerable potential for variation in reliability and objectivity [59]

Clarke, seeing the close relationship of physical education and physical medicine with measures of strength experimented with an aircraft cable control instrument known as a tensiometer [60] With the tensiometer the stress and strain of an aircraft cable can be measured The principle can be applied to human strength by having the subject exert muscle power to put tension on a cable This tension is then measured by the tensiometer, and it varies proportionately to the strength of the subject [61] This instrument is easy to use and comes in different strength cables 5 to 100 pounds 0 to 200 pounds, and 100 to 400 pounds The $\frac{1}{16}$-inch flexible cable attaches by tying to a link chain about 3 feet long, which can be anchored by hook to the wall or floor The other end of the cable connects with the strap that

[59] Lucile Daniels Marian Williams and Catherine Worthington, *Muscle Testing*, Philadelphia, W B Saunders Company 1947

[60] H Harrison Clarke, *Cable tension Strength Tests*, Chicopee, Mass, Brown-Murphy Company, 1953

[61] Cable-tension equipment can be obtained from the Pacific Scientific Company, 1430 Grande Vista Ave, Los Angeles 23, Calif

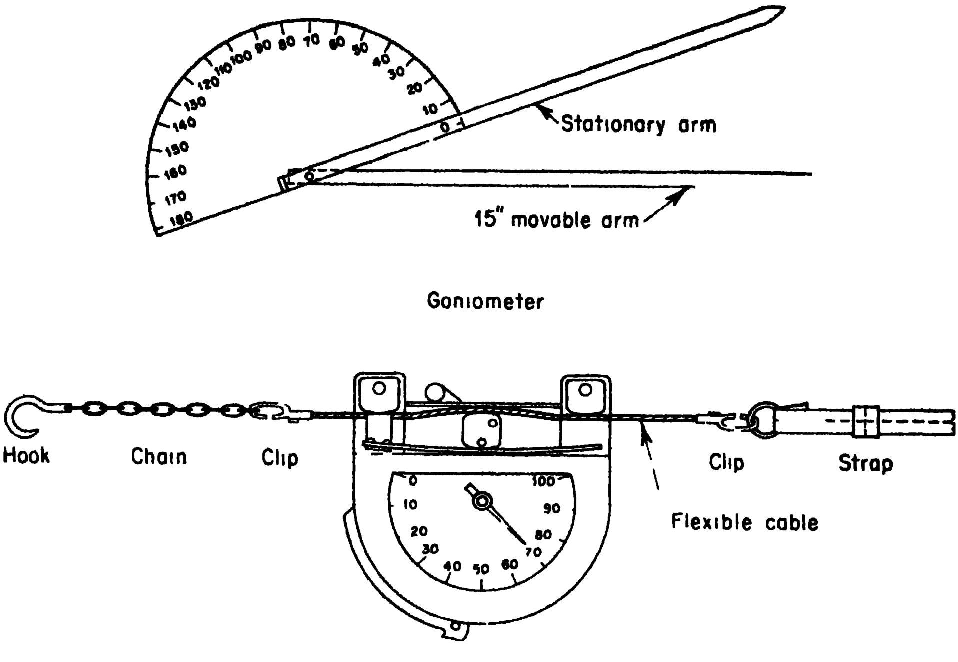


Fig 7 12 Equipment for Cable tension Tests

makes contact with the limb of the subject Thus with the hook attached, the subject pulls at a right angle in any position of the joint

Because tension is measured in a short fixed length of cable, there is no change of cable length to introduce inaccuracy Furthermore, consistent accuracy is maintained because the leverage angle of the subject's limb or position does not change as the load is applied

A goniometer is also used to measure the angle of the joint so that maximum application of force can be obtained A 6-inch protractor with two

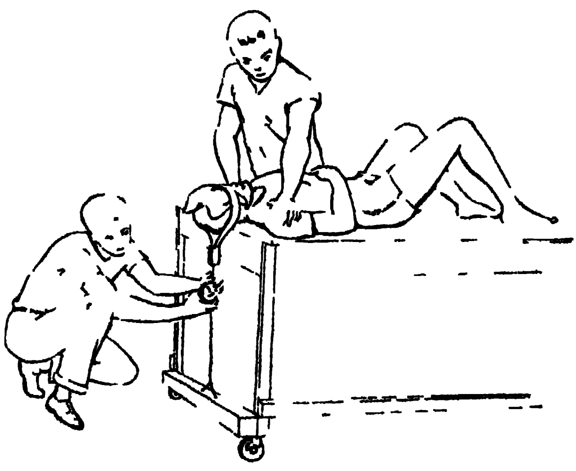


Fig 7-13 Measuring Neck Flexion

15-inch arms is quite satisfactory Additional straps and harnesses for various limbs are also available

In a series of twenty cable-tension tests worked out by Clarke, Bailey, and Shay the test objectivity was quite satisfactory (over 90 for sixteen tests and slightly less for the remaining four) [62] The tests are explained in detail in Clarke's *Cable-tension Strength Tests* As shown in Figure 7-13, it takes two testers to give the test properly One or both testers act to stabilize the subject so he can apply maximum power in an approved manner for any particular measurement A single test example is illustrated in Figure 7-13, where the subject is being tested for neck flexion The position and procedure, as in all cable-tension tests, has been standardized In this particular measurement instructions include the following[63]

Starting position

a Have subject lie in a supine position with hips and knees flexed The elbows should be resting on the table with the hands folded on the chest

b Head should rest on the edge of the table on the occipital lobe

Attachments

a Place the strap over the supra orbital ridge and forehead A pad may be used over the bridge of the nose and forehead for comfort

b Attach pulling assembly below the head

Precautions

a The tester will hold the shoulders flat on the table by pressing down on the subject from above This will also tend to limit any hunching of the shoulders.

b The subject should tuck his chin in when pulling in neck flexion

The future looks bright for cable-tension strength tests, especially in physical medicine, remedial physical education and hospital convalescent programs The tensiometer is accurate and has been used satisfactorily to note individual progress Norms for some of the tests are available for boys ages seven, nine twelve, and fifteen[64]

Clarke, Shay, and Matthews applied the tensiometer to the evaluation of muscle fatigue by setting up a *strength decrement index*[65] This index is the proportionate loss of muscle strength in a given muscle group due to physical exertion It is calculated as follows

$$SDI = \frac{si - sf}{si} \times 100$$

where si = initial strength (taken before exercise)
sf = final strength (taken after exercise)

[62] H Harrison Clarke Theodore L Bailey and Clayton T Shay, "New Objective Strength Tests of Muscle Groups by Cable-tension Methods" *The Research Quarterly*, 23 136–148, May, 1952

[63] *Ibid*, p 143

[64] James Harrison, "The Construction of Cable-tension Strength Test Norms for Boys Seven, Nine, Twelve and Fifteen Years of Age," unpublished master's thesis, University of Oregon, Eugene, Ore, 1958

[65] H Harrison Clarke, Clayton T Shay, and Donald K Matthews, "Strength Decrement Index A New Test of Muscle Fatigue," *Archives of Physical Medicine and Rehabilitation*, 36 376–380, June, 1955

By using the SDI it is possible to note the degree of personal fatigue in a muscle group for a set exercised period. For convalescents with certain disabilities this knowledge can be useful in planning a recovery program. It has potential use in physical education and athletics for estimating how much real effort a person puts into a physical activity.

KRAUS-WEBER TESTS

Having experimented a number of years in the field of physical medicine and rehabilitation, Kraus and his coworkers, at New York University–Bellevue Medical Center, set up a battery of six tests designed to measure the minimum muscular fitness level of children and young adults.[66] Although the tests were somewhat arbitrarily selected, Kraus defends them as tests of *minimum* muscular fitness—a level, below which, individuals are predisposed to orthopedic and emotional difficulties. Specifically the tests were constructed to ascertain "whether or not the individual has sufficient strength and flexibility in the parts of the body upon which demands are made in normal daily living."[67] The experiences of Kraus, Weber, and Hirschland, working with low back pain and other motor disturbances associated with the tension syndrome, seemed to indicate a need for an evaluation of individual muscle strength and flexibility.

Kraus and his associates measured 4,261 American and 2,870 European school children from comparable urban and suburban communities. His test findings showed that 57.9 per cent of the American children failed one or more of the tests compared with 8.7 per cent of the Europeans (Austrians, Italians, Swiss).[68] This caused considerable discussion in physical medicine and physical education circles and was one factor that prompted President Eisenhower to initiate a commission dedicated to the task of raising the physical fitness of American youth. Thus, Kraus had much to do with bringing about a greater interest in the evaluation of physical fitness in the schools and the search for more effective appraisal measures.

Test Equipment Needed: A flat table large enough for a person to recline on and a small pillow are needed.

Testing Procedure: The Kraus-Weber battery of six tests is done on an individual basis. No warm-up is permitted. The tests are as follows:

[66] Hans Kraus and Ruth P. Hirschland, "Minimum Muscular Fitness Tests in School Children," *The Research Quarterly*, 25:178–188, May, 1954.

[67] *Ibid.*, p. 178.

[68] See Hans Kraus and Ruth P. Hirschland, "Muscular Fitness and Health," *Journal of Health, Physical Education and Recreation*, 24:17–24, December, 1953. The difference between these two groups, says Kraus, is the fact that European children do not have the "benefit" of a highly mechanized society; they do not use cars, school buses, elevators, or other laborsaving devices. They must walk everywhere—even to school, frequently a long distance. Their recreation is largely based on the active use of their bodies as compared with America, where recreation as spectators is more common than as participants.

Test 1. A test of abdominal and psoas strength:

Testing position: The person being tested lies in a supine position with the hands behind the neck. The feet are held down on the table by the examiner.

Command: "Keep your hands behind your neck and *try to roll up* into a sitting position."

Precaution: Make certain the subject understands the directions. Permit another try when in doubt. Watch for the "stiff back" sit-up. Possibly the subject didn't realize that the emphasis is on rolling up, or he may have weak abdominal muscles and be doing most of his work with the psoas. Also, watch for a twist of the upper body during the sit-ups.

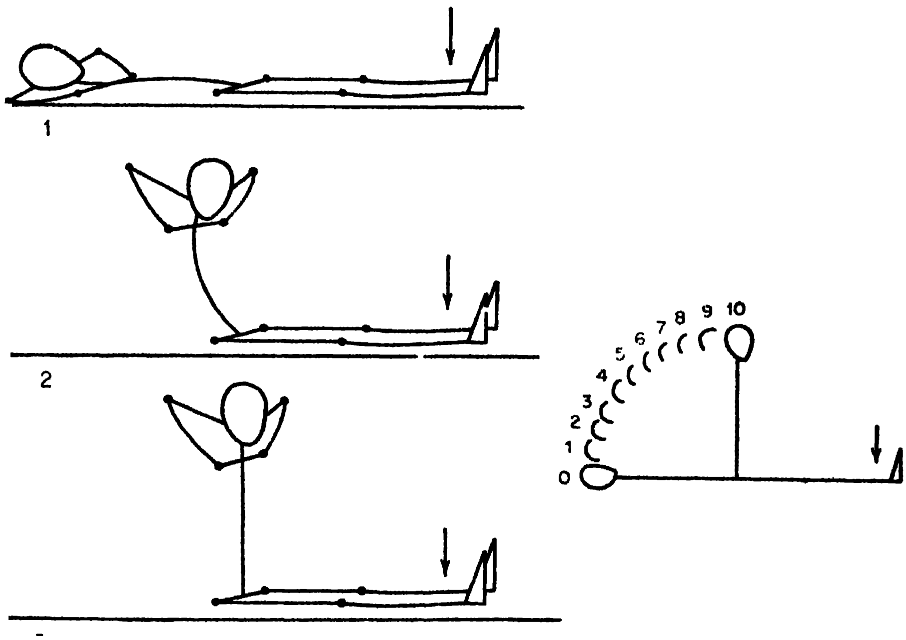


Fig. 7-14. Test 1, Abdominal Muscles Plus Psoas.

Scoring: If the subject cannot raise his shoulders from the table, the score is 0. If, unaided, he is able to reach a sitting position, the score is 10. If he must be helped halfway to the sitting position, the score is 5. The distance from supine to sitting is scored from 0 to 10.

Test 2. A further test of abdominal muscles (minus psoas):

Testing position: The person lies in a supine position, hands behind neck and knees bent. The feet are held down on the table by the examiner.

Command: "Keep your hands behind your neck and *try to roll up* into a sitting position."

Precautions: Precautions are the same as for Test 1. Watch for the "stiff back" sit-up and the tendency to help with one or the other elbow.

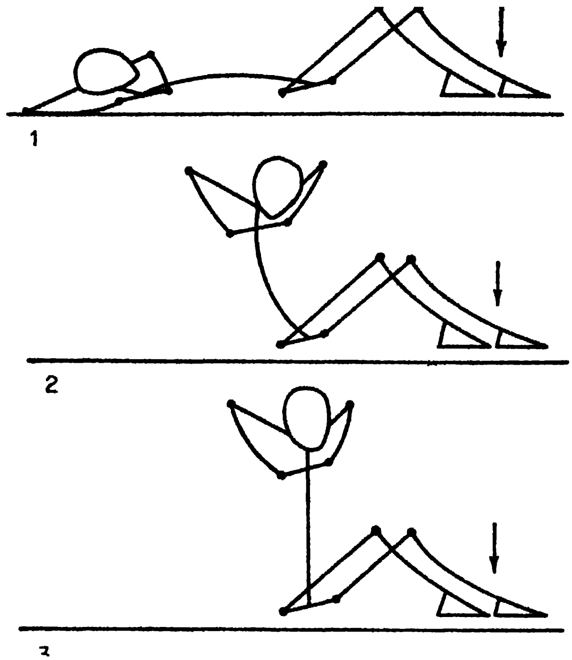


Fig 7-15 Test 2, Abdominal Muscles Minus Psoas

Scoring Scoring is the same as for Test 1

Test 3 A test of psoas and lower abdominal muscles·

Testing position The person lies in a supine position with hands behind neck and legs extended

Command "Keep your knees straight and lift your feet ten inches off the table. Keep them there while I count" The count is ten seconds

Precaution The head and chest must not come up when the feet are raised Very weak abdominal muscles are noted when the back is arched or lordosis is present

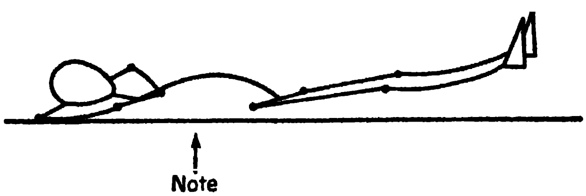


Fig 7-16 Test 3, Psoas and Lower Abdominal Muscles.

Scoring· Holding the count for ten seconds is passing and is scored as 10 Anything less is scored according to the number of seconds the legs are held: 3 for three seconds, 9 for nine seconds, etc

Test 4. A test of strength of upper back muscles:

Testing position: The person lies in a prone position with a pillow under

his abdomen but far enough down to balance the body and give a seesaw effect.

Command: "Roll over on your stomach and lift up the middle so that I can slide this pillow under you. Now I am going to hold down your feet while you put your hands behind your neck and raise up your chest, head, and shoulders. Hold them up while I count." The count is for ten seconds.

Precautions: Watch to see that the chest or elbows are not on the table. Watch also for pronounced muscular development on one side of the spine. If this condition is present, the back should be checked to guard against a scoliosis (lateral curvature of the spine).

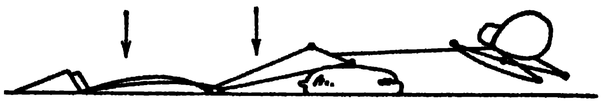


Fig. 7-17. Test 4, Upper Back Muscles.

Scoring: Holding the count for ten seconds is passing and is scored as 10. Anything less is scored according to the number of seconds the back is held up: 4 for four seconds, 5 for five seconds, etc.

Test 5. A test of strength of lower back muscles:

Testing position: The person being tested remains in a prone position over the pillow. He removes his hands from behind his neck and places them on the table and rests his head on them.

Command: "I am going to hold your chest down on the table; try to lift your legs up, but do not bend your knees." It may be necessary to support the legs a little because of the tendency to bend the knees or even to keep the toes on the table. "Now, hold this position while I count." The count is ten.

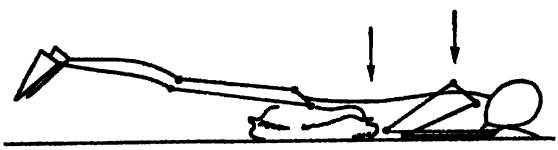


Fig. 7-18. Test 5, Lower Back Muscles.

Scoring: Scoring is the same as for Test 4.

Test 6. A test of the length of back and hamstring muscles:

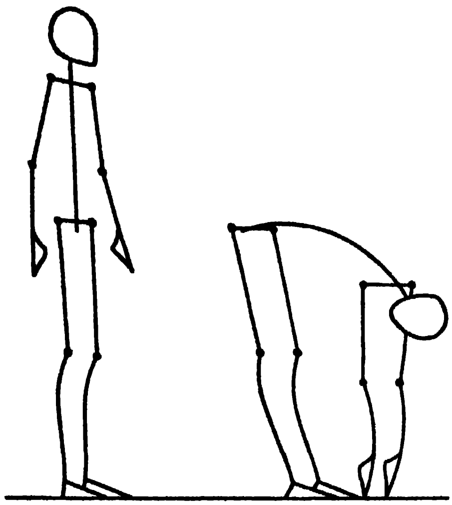
Testing position: The person being tested stands erect in stockings or bare feet with hands at the sides.

Command: "Put your feet together, keep your knees straight; now lean down slowly and see how close you can come to touching the floor with your fingertips. Stay down as far as you can for a count of three. *Do not bounce.*"

Precautions: Bouncing is not permitted. The examiner can control the

performance by holding the knees of the subject being tested in order to prevent bending.

Scoring: The touch (T) is given only when the floor touch is held for three seconds. Less than T is marked by the distance in inches between the



1 2

Fig 7-19 Test 6, Back and Leg Flexibility

floor and the fingertips For instance, a person unable to touch the floor by three inches would be marked —3 inches

Interpretation and Use of Kraus-Weber Tests

These tests are applicable to boys and girls of all ages and have been used with school children and adults alike They are screening tests, which can be administered without involving a great amount of time, equipment, or skill A person failing one part of the test is considered to have failed the test as a whole

A fair amount of discussion has taken place over the validity of the Kraus-Weber battery as a measure of physical fitness The critics insist that the test battery was arbitrarily selected for use with persons suffering from low back pain There are no norms Also, some persons have no trouble with the first five test items but fail Test 6 (flexibility). Kraus contends, however, that the tests are not meant to measure degrees of physical fitness as a more elaborate performance test might do; instead, the six tests together constitute only a minimum muscular fitness test, which provides a minimum standard for children Furthermore, the tests are self-validating and do not need norms because the subject is only being compared with himself. To fail part of the test, says Kraus, indicates that the subject is below standard and predisposed toward orthopedic or emotional difficulties

In an Indiana city, 1,456 elementary school children were tested with the Kraus-Weber Tests. The tests were highly reliable as evidenced by coefficients of .95 in all cases.[69] The Indiana group was found to be somewhat superior to the Kraus group in all failure comparisons; girls were superior to boys at all levels in passing the test and in flexibility; for both sexes there was a decided decrease in strength item failure as age increased, so much that by eleven years of age the percentage of failure on any test item was less than 8 per cent and agreed closely with results found for European children. In the same study it was found that not over six children in the entire study population failed on each of the back strength items, which suggests that these items do not discriminate well.[70] Noguchi made the same suggestion after his research with the test on 6,549 Japanese school children, where he found no failures of Tests 3, 4, and 5.[71] His studies coincided with those of Kraus in that the percentage of Japanese failures increased between twelve and thirteen years of age. Only 3.3 per cent of the Japanese children failed the test of flexibility.

In an effort to make the Kraus-Weber scores more meaningful, Buxton studied the tests in terms of age and sex groups and found that strength and flexibility differ with age and sex enough to encourage the development of standards.[72]

The question of flexibility and its relationship to physical fitness seems to be a stumbling point in the Kraus-Weber battery. Noguchi questioned whether flexibility related to body arm-trunk-leg proportions. Gurewitch and O'Neil found that flexibility varied with age.[73] Kendall and Kendall also pointed this out.[74] Fox and Atwood, working with Iowa City children, urged caution in interpreting the results of the Kraus-Weber toe-touch test.[75] Broer and Galles, in a study involving college women, showed that trunk plus arm length/leg length is not an important ratio in the performance of the toe-touch test for those persons with average body builds.[76] Among ex-

[69] Marjorie Phillips, Carolyn Bookwalter, Charlotte Denman, Janet McAuley, Hilda Sherwin, Dean Summers, and Helen Yeakel, "Analysis of Results from the Kraus-Weber Test of Minimum Muscular Fitness in Children," *The Research Quarterly,* 26:314-323, October, 1955.

[70] *Ibid.*, p. 322.

[71] Yoshiyoki Noguchi, "Fitness Testing of Japanese Children," *Journal of Health, Physical Education and Recreation,* 27:20, October, 1956.

[72] Dorothy Buxton, "Extension of the Kraus-Weber Test," *The Research Quarterly,* 28:210-217, October, 1957.

[73] A. D. Gurewitch and Margaret A. O'Neil, "Flexibility of Healthy Children," *Archives of Physical Therapy,* 27:216-221, April, 1944.

[74] Henry O. Kendall and Florence P. Kendall, "Normal Flexibility According to Age Groups," *Journal of Bone and Joint Surgery,* 30-A:690-694, 1948.

[75] Margaret G. Fox and Janet Atwood, "Results of Testing Iowa School Children for Health and Fitness," *Journal of Health, Physical Education and Recreation,* 26:20-22, September, 1955.

[76] Marion R. Broer and Naomi R. G. Galles, "Importance of Relationship between Various Body Measurements in Performance of the Toe-touch Test," *The Research Quarterly,* 29:253-263, October, 1958.

treme types, however, the relationship is more significant. Matthews, Shaw, and Woods also discounted as insignificant the body length and lower limb length measurements as they relate to flexibility in the toe-touch test.[77]

As in all tests, the *use* to which they are put is the important item for consideration The use of the Kraus-Weber battery at the elementary school level for screening boys and girls in need of remedial physical education is commendable. Although the flexibility item accounts for a great number of failures, the usefulness of the test is not seriously affected Kirchner and Glines found that it could be employed to screen children in Oregon even though flexibility accounted for the greatest number of failures[78] Shaffer, working with 1,400 junior high school girls in two schools with different physical education programs, found that Kraus-Weber Tests failure is positively correlated with intelligence, that is, as intelligence increased, test failures decreased[79] Also, as the age level of the girls increased, the percentage of test failures increased Furthermore, it was found that a program of exercises based on physiological needs produces rapid gains in strength and flexibility, for after one semester of a program including conditioning exercises, these girls, most of whom had had no physical education before the experiment, matched the European children's low rate (8 7 per cent) of failure in the Kraus-Weber Tests Moreover, Shaffer points out that with these conditioning exercises, "requiring less time than is necessary for the classes to take showers, done twice each week for part of two semesters, . . . junior high school students have less than 5 per cent Kraus-Weber Test failure The improvement in appearance and well-being of the girls is far more rewarding than any test results"[80]

ENDURANCE TESTS

It is often difficult to separate muscular endurance from measures involving muscular strength, power, speed, and flexibility In fact, endurance is certainly a prime ingredient in physical fitness and general motor ability It is characterized by capacity for sustained activity and is illustrated by running, dipping, chinning, and swimming

In a factor-analysis study of twelve athletic events, administered to 400 well-conditioned soldiers, McCloy found four distinct factors: circulo-respiratory endurance, velocity, muscular endurance, and mesomorphic body

[77] Donald K Matthews, Virginia Shaw, and John B Woods, "Hip Flexibility of Elementary School Boys as Related to Body Segments," *The Research Quarterly*, 30 297–302, October, 1959

[78] Glenn Kirchner and Don Glines, "Comparative Analysis of Eugene, Oregon Elementary School Children Using the Kraus-Weber Test of Minimum Muscular Fitness," *The Research Quarterly*, 28:16 25, March, 1957.

[79] Gertrude Shaffer, "Variables Affecting Kraus-Weber Failures among Junior High School Girls," *The Research Quarterly*, 30 75 86, March, 1959.

[80] Gertrude Shaffer, personal correspondence, December, 1959

build[81] The weighting of variable 6 (see the following chart) of 8835 is almost as high as that for the whole endurance combination, variable B (8964).

Rotated Factor Loadings

	Event	Circulorespiratory endurance I	Velocity or muscular contraction speed II	Muscular endurance III	Mesomorphic build IV	Communality h
1	Shot put	3085	4414	2501	1351	6089
2	Three broad jumps	2765	5993	3526	0789	7525
3	Shuttle run	248	6720	0423	1475	7327
4	6-sec sprint	0941	6363	5045	0933	8226
5	300 yd run	7948	4417	1243	1001	9231
6	300/6-sec run	8835	0652	1953	1027	9130
7	75 yd Pickaback run (with man on back)	6139	3264	3304	5794	9634
8	Pick/6 sec run (with man on back equated with 6 sec run)	6542	1335	1660	6302	9330
9	20 sec squat thrusts	4893	1318	4531	0327	6807
10	Chinning or pull ups	3977	0861	4263	0349	5908
11	Push ups	5525	0366	5746	0184	7982
12	Sit ups	4674	0279	5821	0442	7484
A	Muscular endurance events composite	5835	1671	7178	0103	9401
B	Circulorespiratory endurance events composite	8964	2695	1152	2790	9835
C	Speed events composite	2482	8693	3441	0894	9956
D	Total points all tests	7221	5119	4427	0136	9897

SOURCE C H McCloy A Factor Analysis of Tests of Endurance *The Research Quarterly* 27 213 216 May 1956

McCloy's Endurance Ratio

The proportion of times between short and long runs is considered an endurance ratio A distance run is divided by the time for a sprint In McCloy's test the subject runs 300 yards This can be done on a track or over a 60-yard course Later the subject performs a six-second sprint, with the distance run being recorded to the nearest foot and converted to units of time[82] Scoring is as follows

$$\frac{\text{Time for 300-yd run}}{\text{Distance for 6-sec run}}$$

[81] Charles H McCloy "A Factor Analysis of Tests of Endurance," *The Research Quarterly,* 27 213 216, May 1956

[82] Charles H McCloy and Norma D Young, *op cit*, pp 183–189

Endurance Running

Ordinary all-out straight running over a 200- , 300- , or 400-yard course is useful in estimating the endurance level of school children. The 300-yard shuttle run used by some of the Armed Forces during World War II was a satisfactory test because it was just long enough to eliminate the factor of speed and short enough to eliminate the knowledge of pace, an item so necessary in longer events such as the mile and two-mile run. In endurance running, however, the subject must be properly motivated so he will not ease off but will strive to do his best. As the endurance event is continued and rapid respiration and great fatigue set in, the will of the boy or girl to continue is severely taxed. Few persons want to push themselves to exhaustion. Without adequate motivation, therefore, reliability of the endurance measure is apt to be exceedingly low.

The Drop-off Index

In the drop-off index a person's time for an endurance event is recorded in time intervals as endurance begins to wane and speed is reduced. It is the rate of deceleration. It can be figured by comparing the time of the first lap (in a swimming pool or on running track) with the last lap. As the endurance improves, the drop-off between first and last lap times will decrease. Such a measure can aid the coach to determine whether his teams have improved in their ability to forestall fatigue. Cureton, for example, worked out a drop-off index for comparing swimmers.[53] The subject swims the 100 yards in a 60-foot pool. The time for the first lap (60 feet) is noted and compared with the slower time of the final lap. For a man running the 1,000-yard run and the 100-yard dash the drop-off index would be figured by the formula Drop-off index = 1,000-yd run time $-10 \times$ 100-yd run time.

SELECTED REFERENCES

Carpenter, Aileen: "A Critical Study of the Factors Determining Effective Strength Tests for Women," *The Research Quarterly*, 9:3 32, December, 1938.

Clarke, H. Harrison: "Developing Physical Fitness Programs in Small Schools," *Education*, 60:472 477, April, 1940.

———: "Oregon Pilot Physical Fitness Project," *The Physical Educator*, 14:55–58, May, 1957.

———: "Physical Fitness Benefits: A Summary of Research," *Education*, 78:460–466, April, 1958.

Cousins, George: "Effect of Trained and Untrained Testers upon Administration of Grip Strength Tests," *The Research Quarterly*, 26:273–278, October, 1955.

Cozens, Frederick W.: "Strength Tests as Measures of General Athletic Ability in College Men," *The Research Quarterly*, 11:45-52, March, 1940.

[53] Thomas K. Cureton, "A Test for Endurance in Speed Swimming," *The Research Quarterly* (Supplement), 6:106–112, May, 1935.

Hall, D. M.: "Selection and Standardization of Strength Tests for 4-H Members," *The Research Quarterly,* 27:285–295, October, 1956.

Hines, Thomas H.: "A New Emphysis upon Health," *Journal of Health and Physical Education,* 10:20, January, 1939.

Hunsicker, Paul, and R. Donnelly: "Instruments to Measure Strength," *The Research Quarterly,* 26:408–420, December, 1955.

Kelliher, M. S.: "A Report on the Kraus-Weber Test in East Pakistan," *The Research Quarterly,* 31:34–42, March, 1960.

Kraus, Hans, and Ruth P. Hirschland: "Muscular Fitness and Health," *Journal of Health, Physical Education and Recreation,* 24:17–24, December, 1953.

Matthews, Donald K.: *Measurement in Physical Education,* Philadelphia, W. B. Saunders Company, 1958, pp. 48–56.

McCloy, Charles H.: *Measurement of Athletic Power,* New York, A. S. Barnes and Company, 1932.

Methenny, Eleanor: "The Present Status of Strength Testing for Children of Elementary School and Pre-school Age," *The Research Quarterly,* 12:115–130, March, 1941.

Rogers, Frederick Rand: *Physical Capacity Tests,* New York, A. S. Barnes and Company, 1930.

———: "A Review of Recent Strength Testing Literature," *Journal of Health and Physical Education,* 5:62–65, March, 1934.

———: "The PFI Analogue and Complement of I.Q.," *School and Society,* 50:546–547, September, 1939.

———: "The Evolution of PFI Tests and Programs," *Education,* 60:527–540, April, 1940.

Scott, M. Gladys, and Helen Matthews: "A Study of Fatigue Effects Induced by a Efficiency Test for College Women," *The Research Quarterly,* 20:134–138, May, 1949.

Wells, Philip V.: "Emotion in Fitness Tests," *The Research Quarterly,* 26:258–262, October, 1955.

Willgoose, Carl E.: "Use of Strength Tests in Team Equalization," *The Physical Educator,* 17:704-712, March, 1949.

———: "The Relationship of Muscular Strength to Motor Coordination in the Adolescent Period," *Journal of Educational Research,* 64:138–142, October, 1950.

Wright, Verna: "Factors Influencing Diurnal Variations of Strength of Grip," *The Research Quarterly,* 30:110–116, March, 1959.

Zankel, Harry T., et al.: "Physical Fitness Index Studies (PFI) in Hospitalized Diabetic Patients," *Archives of Physical Medicine and Rehabilitation,* 38:250–258, April, 1957.

———, and James N. Field: "Physical Fitness Index in Psychiatric Patients," *Journal of the Association for Physical and Mental Rehabilitation,* 40:395–401, March–April, 1959.

CHAPTER 8

Measures of Physical Fitness
Motor Performance Tests

The two previous chapters have dealt basically with two categories of physical fitness tests cardiovascular tests and measures of muscular strength and endurance Related to these tests is a wide variety of fitness measures that may be termed essentially *motor fitness* or *motor performance tests* [1] These are tests that require the subject to do more than exert strength or demonstrate endurance as such Motor fitness is a *qualitative* performance involving the physical fitness elements of heart-lung efficiency, muscular strength, and muscular endurance in combination with power, speed agility, flexibility, and balance [2] The subject being measured frequently performs in a number of fundamental activities such as running, jumping, and throwing The element of skill is held to a minimum in an effort to get at the physical capacity that supports performance

The Variety of Tests

There are a great number of measures designed to appraise some aspect of physical fitness No attempt will be made to cover them all in this chapter Only those better known measures that appear to meet the criteria of proper tests will be included

Since President Eisenhower's first conference pointing to the physical fitness needs of American youth, a great number of tests have appeared These are being used to a limited degree by public and private schools, youth organizations, the Armed Forces, and adult groups in the community Several of these tests have been carefully developed under the auspices of state education departments, of which California and New York are good examples, and, in some instances, under the guidance of state associations

[1] See Chap 6 for a discussion of physical fitness, together with definitions of physical fitness, motor fitness, cardiovascular fitness, muscular strength, and muscular endurance

[2] Cureton defines motor fitness as "a limited phase of motor ability which emphasizes capacity for vigorous work or athletic effort", see "An Inventory and Screen Test of Motor Fitness for High School and College Men," *The Physical Educator*, January, 1943

for health, physical education, and recreation The AAHPER Youth Fitness Test was developed at the national level by the American Association for Health, Physical Education, and Recreation

AAHPER YOUTH FITNESS TEST

In little over one year the fitness test prepared for the American Association for Health Physical Education, and Recreation was given to over 14,500 000 children and young persons as a part of a test program in fifty states and sixteen foreign countries [1] This test is part of a kit of materials designed to promote higher levels of physical fitness for boys and girls aged ten to seventeen in schools camps, recreation departments, youth agencies, and institutions (Operation Fitness USA) An award is made in connection with successful completion of the test at senior high, junior high, and elementary school levels

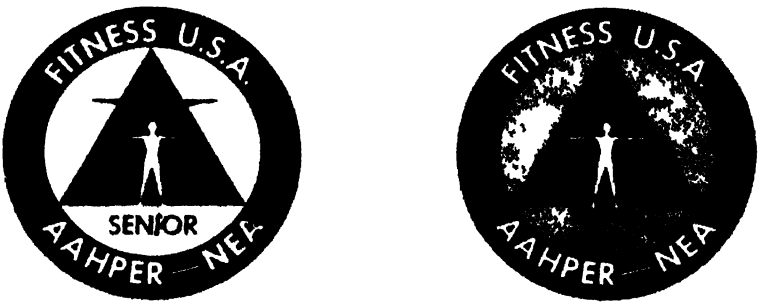


Fig 8 1 Senior High School Award Operation Fitness USA (AAHPER Youth Fitness Test) Left award for senior high school boys and girls grades ten eleven and twelve Standard emblem for fiftieth percentile Right merit emblem for eightieth percentile

In this test a committee selected seven test items pull ups sit-ups, 40-yard shuttle run standing broad jump 50-yard dash, softball throw for distance, and 600-yard run-walk Three aquatic tests are also included Percentile norms are available based both on age and on the Neilson and Cozens (California) Classification Index Data for the norms were obtained from a sample of 8 500 boys and girls in grades five through twelve from schools throughout the United States

Test Equipment The following items are needed metal horizontal bar adjustable to various heights floor mat, two blocks of wood, 2 by 2 by 4 inches, stop watch, softball, tape measure

Test Procedure Record height and weight in gym suit and in stocking

[1] American Association for Health Physical Education and Recreation *Progress Report during the First Few Months of Action, January to December 1959*, Washington, December, 1959 See also *AAHPER Youth Fitness Test Manual*, American Association for Health Physical Education, and Recreation, Washington, 1958.

feet. Include the pull-up test (or modified for girls), the sit-up, shuttle run, and 50-yard dash in the first test period. Give the standing broad jump, the softball throw for distance, and the 600-yard run-walk in the second test period. The aquatic tests are given whenever suitable swimming facilities are available.

Pull-up: After assuming the hanging position the subject raises his body by his arms until the chin can be placed over the bar. The body is then lowered to a full hang as in the starting position. This is repeated as many times as possible. There must not be any swinging or snapping movements during the test.

Pull-up (modified): The girl or boy, in the fifth and sixth grade, grasps the bar, which is placed at shoulder height. A reversed grip is used. The legs are extended under the bar and the chest is held at the bar. (In the Connecticut modification the bar is placed at a fixed height of 3 feet so that it does not have to be adjusted for each subject.) The body is held in a straight line, and the heels are braced on a mat to prevent slipping. The subject lowers the body until the arms are fully extended and returns by pulling with the arms until the chest touches the bar. The exercise is repeated as many times as is possible with a limitation of a maximum number of forty repetitions.

Sit-up: The subject lies flat on a mat with the legs straight and the feet about 2 feet apart. The hands are placed on the back of the neck with fingertips touching or fingers clasped. Elbows are flat on the mat. A partner holds the ankles, keeping the heels in contact with the mat at all times. When ready, the subject sits up touching the left elbow to the right knee. This is repeated, alternating sides. Each complete sit-up counts one point. No score is given if the fingers slip off the neck, or the knees are bent, or the subject pushes off the floor from an elbow. The maximum limit of sit-ups is 50 for girls, grades five to twelve, and 100 for boys, grades five to twelve.

Shuttle run: Two parallel lines are drawn on the floor 30 feet apart. The blocks of wood are placed just beyond the second line, about a foot apart. Space is needed behind both lines so that pupils can overrun the lines without danger.

The subject stands behind the first line. On the signal "Are you ready? Go!" the subject runs to the second line, picks up one block of wood, runs back to the first line and *places* (not throws) the block behind the line. He then runs back to the second line, picks up the second block of wood, and runs back to the first line, finishing as he crosses the line with the block in his hand. (If pupils start alternately from first and second lines, it will be unnecessary to move blocks between runners, and time will be saved.)

Two trials are permitted. Time is recorded in seconds and tenths of seconds from the starting signal until the runner crosses the line on the second round trip. Record the better time of the two trials.

Fifty-yard dash: After a short warm-up the subject takes a position behind

the take-off line. The starter uses the commands "Get on your mark; Get set; Go!" The word "go" is accompanied with a downward sweep of the starter's arm as a signal to the timer.

More than one subject may run at a time if additional watches are available. Only one trial is given per class period. The score is recorded in seconds and tenths of seconds from the starting signal to the moment the subject crosses the finish line.

Softball throw for distance: Two lines are drawn parallel and 6 feet apart, within which limits the throw is made. The line nearest the area toward which the ball is thrown is called the restraining line. To measure quickly the distance of the thrown ball, additional lines can be drawn parallel to the restraining line and 10 yards apart as far down the field as most balls will land.

After a little warm-up the subject throws the ball as far as possible from behind the restraining line. One or more steps may be taken, provided the subject remains behind the line. An overhand throw must be used.

Three throws are made in succession, and the best of these is recorded. Measurement is made from the restraining line to the point where the ball first lands. Distance is recorded to the nearest foot.

Standing broad jump: The subject stands with the feet several inches apart and with the toes just back of the take-off line on the floor or mat. The take-off is made from both feet and the subject jumps forward as far as possible, landing on both feet.

Three fair trials are permitted, and the best of the three is recorded. Measurement of the jump is made from the nearest imprint (including any imprint by hands or body) made by the jumper in landing to the take-off point. Distance is recorded in feet and inches.

Six-hundred-yard run-walk: From a standing start and from the "go" signal, the subject starts running the 600-yard distance. The running may be interspersed with periods of walking. Encourage the subject to do his very best. The event is timed to the nearest second, from the starting signal until the required distance is covered.

Aquatic test: These tests do not figure in the norm table and in many instances are omitted for lack of swimming facilities.

Test 1. The subject starts in water and swims a distance of 15 feet, using a stroke or strokes of his own choice. Test is scored as "pass" or "fail."

Test 2. The subject jumps into water over his head. He is then required to swim 15 yards; turn around; swim back half the distance; turn on back and rest for one-half minute; turn over to front; and swim to starting point. Test is scored as "pass" or "fail."

Test 3. On the "go" signal the subject dives into the water and swims 100 yards. Time is recorded to the nearest second.

Scoring: The separate test item scores are compared with percentile scores on a norm table. A norm table for each test item and each sex appears

in the *AAHPER Youth Fitness Test Manual.* The norms are based on the Neilson and Cozens (California) Classification Index, a copy of which appears in Chapter 14 Percentile norms for college men appear in the Appendix.

Test Interpretation: As in all tests, the *use* to which they are put influences their worth The whole national movement (Operation Fitness, USA) serves a long range purpose when test results are studied and individual pupil programs are modified accordingly Follow-up work, with proper personal motivation to improve, is important The award program has value if the awards are handled with proper controls and respect The American Red Cross, Boy Scouts of America, and Girl Scouts of America have successfully employed awards for many decades to improve and give meaning to their program

In the AAHPER Youth Fitness Test the stress should be on minimum norms, for they may appear too easy for children to meet in certain areas of the United States Well-constructed local norms for this test may prove useful

Because the norms are based on percentile scores the scale values are close together around the median (fiftieth percentile) and spread out at the extremities of the distribution For example, five-scale-point increments for the softball throw for girls Classification A, are as follows. from ninety-fifth to hundredth, 27 feet, ninetieth to ninety-fifth, 5 feet, fiftieth, to fifty-fifth, 1 foot; zero to fifth, 8 feet The effect of such a scale is to encourage mediocrity, because small gains in performance around the average result in rapid advancement on the scale For the very good and very poor performers, however, great gains in performance result in slow improvement on the scale Some form of the standard-scoring process, such as the T scale, or 6-sigma scale tends to avoid these discrepancies, as the entire distribution is divided into equal increments The use of the classification index (age-height-weight combination) is a good choice, this process is also used effectively in the Indiana and the California motor fitness tests

NEW YORK STATE PHYSICAL FITNESS TEST

Since early in the 1920s the New York State Education Department has been advocating some form of physical fitness test Hundreds of school systems have successfully employed the Physical Fitness Index and similar tests The New York State Physical Fitness Test is "designed to provide schools with a convenient instrument for periodic evaluation of status and progress in physical fitness of boys and girls in grades four through twelve."[4] Moreover, the test can be used as a motivating device for the pupil, as a diag-

[4] *The New York State Physical Fitness Test: A Manual for Teachers of Physical Education,* Albany, N Y, Division of Health, Physical Education, and Recreation, New York State Education Department, 1958

nostic aid for the teacher, and as a source of objective information for the parent.

There are seven different test items used to measure seven components of physical fitness. These components were selected from a list of sixty-nine physical fitness elements, which fell into three categories: (1) medical or physiological function, (2) anthropometric condition pertaining to physique or appearance, and (3) physical or motor function. During early experimentation a number of components were eliminated and others kept in by the criteria of leadership, time, scheduling, and equipment and facilities. The final selection of test items was made on the basis of their validity, reliability, and administrative feasibility as indicated by pretest results.

A brief description of the test items used to measure each component is given below:

Posture: Posture is evaluated by means of a *posture rating chart* The subject is compared with the figure drawings illustrating posture for 13 different body segments Each body segment is scored as 5, 3, or 1, making a possible range of scores from 13 to 65

Accuracy. The *target throw* is used to measure accuracy The subject makes twenty throws with a softball at a circular target and is scored on the number of hits

Strength: Pull-ups (chinning) for boys and *modified pull-ups* for girls are used to measure strength in the seventh grade and above. *Modified push-ups* against a bench are used in grades four through six for both boys and girls. The score is the number of pull-ups or push-ups completed

Agility: The side step is used to measure agility. Starting from a center line, the subject side-steps alternately left and right between two lines 8 feet apart He is scored on the number of lines crossed in ten seconds.

Speed: The *50-yard dash* is used to measure speed. The score is the amount of time to the nearest half second.

Balance: The *squat stand* is used to measure balance The subject squats with elbows against the inner knee surfaces and leans forward until the feet are raised just off the floor. He is scored on the number of seconds he holds his balance

Endurance: The *treadmill* is used to measure endurance. The subject starts from a modified front-leaning rest position, with hands on mat shoulder width apart, one knee flexed, and the other extended. He then performs by exchanging the positions of his feet. The score is the number of leg changes in thirty seconds for boys and girls in grades four, five, and six and for girls in grades seven through twelve; in one minute for boys in grades seven through twelve.

The test is scored by recording the raw scores on each test item and relating these to the achievement level score found on the norm tables The seven achievement level scores are totaled and compared with the total

physical fitness achievement norm for the sex and grade of the subject. The example of the boy in Figure 8-2 indicates that he did five pull-ups with 3 achievement-level points. These points, together with the points from the other six items, total 50, a figure that rates a 9 (ninetieth percentile) for the thirteen-year-old boy

The norms were obtained by administering the test to 12,626 boys and girls in twenty-six different school systems in New York State. A profile chart for sending home to the parents and a cumulative record chart for the instructor are also in use as part of the evaluation program The norm tables and the posture rating chart have been omitted from this text only because they are elaborate and space-consuming Copies of the New York State Physical Fitness Test manual are available on request from the New York State Education Department, Albany 1, New York.

NEW YORK STATE PHYSICAL FITNESS TEST
PUPIL SCORE CARD
Name O'Toole James M.
Last First Middle

GRADE	8	
DATE	Mo. 10 19 61	
AGE	Yrs. 13 Mos. 3	
HEIGHT	Ft. 5 In. 3	
WEIGHT	Lbs. 110	
COMPONENT	raw score	achievem't level
1. Posture	63	8
2. Accuracy	11	8
3. Strength	5	3
4. Agility	16	7
5. Speed	7.0	8
6. Balance	55	10
7. Endurance	51	6
Total Physical Fitness		SUM = 50
		9

EXAMINER CEW DATE 10/25/61

Fig. 8-2 Score Card for New York State Physical Fitness Test.

CALIFORNIA PHYSICAL PERFORMANCE TEST

In California, physical fitness testing began as far back as 1918, under the able direction of Clark Hetherington. During World War II the *California Physical Fitness Pentathlon* was used in California to measure aspects of physical fitness for boys from junior high school to college[5] Later, in an effort to improve on this test, the *California Physical Performance Test* was developed.[6]

In many ways the test is similar to the AAHPER Physical Fitness Test. In fact, the California research had some bearing on the selection of test items for the AAHPER test. Events are grouped into five groups for boys and girls according to grade level. There are "preferred," "alternate," and "optional" tests, according to pupil interests, facilities, and equipment

Figure 8-3 illustrates the similarity of the California and AAHPER tests.

[5] *California Physical Fitness Pentathlon,* Bulletin of the California State Department of Education, Sacramento, Calif., vol. XI, no. 8, November, 1942.

[6] *California Phyical Performance Test,* Bureau of Health Education, Physical Education, and Recreation, California State Department of Education, February, 1958.

Each event in the test is administered once each semester, twice if feasible. Practice and extended instruction is encouraged so pupils will develop endurance and skill. No more than three tests are given in any one class period. Raw test scores are related to the norm tables according to the California

	Events	Boys events, grades			Girls events, grades		
		5–6	7–9	10–12	5–6	7–9	10–12
California Physical Performance Test	Group I						
	Standing broad jump*	P	P	P	P	P	P
	Jump and reach	A	A	A	A	A	A
	Group II	one test only					
	Pull-up (modified)						
	Pull-up*		P	P			
	Push-up		A	A			
	Push-up (knee)				O	O	O
	Group III						
	Sit-up*	O	O	O	O	O	O
	Group IV						
	50-yd dash*	O	P	P	O	P	
	75-yd dash		optional (ninth grade only)				
	Shuttle race*						
	Group V						
	Throw for distance:						
	Softball*	O	P	P	O	P	P
	Basketball or soccer ball		A	A		A	A
AAHPER test (items not in California test)	Run and walk, 600 yd		optional ninth grade only				optional
	Swim, 15 ft		all optional				
	Jump and swim, 30 yd		all optional				
	100-yd free-style swim		all optional				

* California Events selected for AAHPER test. P = Preferred; A = Alternate; O = One test only.

Fig. 8.3. Comparison of California Physical Performance Test with AAHPER Test.

plan of classification, which involves assigning exponents to the various ages, heights, and weights and adding the exponents for the individual to determine his classification. See reference to Neilson and Cozens (California) Classification Index, Chapter 14. The classification chart works as follows:

Example

	Exponent
Elementary school boy or girl	
Height, 57 inches	6
Age, 13 years 2 months	7
Weight, 102 pounds	9
Sum of exponents	22
Pupil is in class	D

Test directions already given for the AAHPER Physical Fitness Test can be followed and raw scores compared by classification index, sex, and grade to the California test by referring to Table A-1, Appendix A.

UNIVERSITY OF MARYLAND MOTOR FITNESS TEST

Research has indicated that such activities as pull-ups, push-ups, sit-ups, and a jumping event are valuable items in fitness tests It is not uncommon, therefore, to discover a number of tests having the same events The motor fitness test at the University of Maryland consists of pull-ups, the vertical jump, fifteen squat-thrusts, and sit-ups[7] Chinning and the vertical jump are done in the usual way The squat-thrusts are performed as fast as the subject can move The event is scored by the number of seconds it takes to accomplish the fifteen correct squat-thrusts Test reliability and objectivity is high, and raw scores can be transposed into T scores

UNIVERSITY OF FLORIDA PHYSICAL FITNESS TEST

Another test requiring little equipment and relatively easy to administer to college men is the physical fitness test used at the University of Florida.[8] There are four events: (1) push-ups are performed on the ends of the parallel bars and scored according to number properly performed, (2) sit-ups have a two-minute time limit and are scored according to the number performed, (3) the 100-yard dash is scored according to seconds and tenths of seconds; and (4) the 300-yard run is scored to the nearest second The raw scores for each event are related for standardization to T-score tables.

YALE UNIVERSITY PHYSICAL FITNESS TEST

At Yale University fitness testing is an established part of the freshman required program of physical education.[9] The test, started during the years

[7] Test in mimeograph form can be obtained from the Required Physical Education Department, College of Physical Education, Recreation, and Health, University of Maryland, College Park, Md

[8] Test copies together with scoring tables are available from the Department of Required Physical Education for Men, College of Physical Education and Health, University of Florida, Gainesville, Fla

[9] T. Erwin Blesh and Alfred E. Scholz, "Ten-year Survey of Physical Fitness Tests at Yale University," *The Research Quarterly,* 28:321–326, December, 1957.

1942 to 1945, consists of six items They are listed below together with the standard performances calculated from the records of approximately ten thousand freshmen men (1947 1956)

1 Pull-ups (standard, eight times)

2 Push-ups (standard, twenty-five times), done from the floor

3 Sit-ups (standard, fifty times)

4 Fence-vault (standard, 4 feet 6 inches) a horizontal bar vault executed without touching either hips or legs to the bar Five or six trials are permitted

5 Standing broad jump (standard, 86 inches) executed on a mat Three trials are permitted

6 Vertical jump (standard, 18 inches)

Students failing to meet the standard for these tests are assigned to the remedial physical education program where nearly 80 per cent of the persons who fail one or more tests in the initial testing period pass the six tests during the twelve-week program of three half-hour periods per week This is an excellent example of how physical fitness tests can be used

UNIVERSITY OF ILLINOIS MOTOR FITNESS TESTS

In his study of motor fitness Cureton analyzed a number of performance skills in terms of the fitness elements of balance flexibility, agility, strength, power and endurance [10] It is Cureton's premise that these elements form a "mosaic of fundamental abilities contributing to gross motor coordination Cureton makes a sound point that it is not the purpose of motor fitness tests to test highly specialized skills but the purpose is "to test the fundamental motor fitness characteristics which *underlie* gross performance in these activities "[11]

Two tests of motor fitness were developed at the University of Illinois, in 1941, and used extensively over the years at that institution The first test represents the fourteen most valid items as correlated with a thirty-item motor fitness inventory (.87) The second test somewhat similar, is an eighteen-item test for adult men and women designed to eliminate all apparatus and can be given on a bare gymnasium floor or outdoor field In this test the subject is checked off as having passed or failed the item The number of passed items is totaled and the classification obtained A subject passing fifteen of the eighteen tests would rate 83 3 per cent for a classification of Good (see Figure 18-4)

A seven-item *short motor fitness test* was also developed at the University of Illinois, but validity is limited [12] Also, a *motor fitness screen test for*

[10] Thomas K Cureton, *Physical Fitness Appraisal and Guidance*, St Louis The C V Mosby Company, 1947, p 393

[11] *Ibid*, p 394

[12] Thomas K Cureton, Lyle Welser, and W J Huffman, "A Short Screen Test for Predicting Motor Fitness," *The Research Quarterly*, 16 106–119, May, 1945

Item	Check result*		Ring result		Class
	Pass	Fail	Passed	%	
Balance: 1. *Diver's stance*, on toes, eyes closed, 20 sec			18	100	Superior
2. *Squat stand*, 10-sec balance			17	94.4	
3. *Dizziness recovery*, walk 10-ft line, 5 sec after 10 turns around finger on floor			16	88.9	
Flexibility: 4. *Floor touch*, knees straight (women touch palms flat)			15	83.3	Good
5. *Trunk flexion, forward*, sitting position, knees held down, forehead lowered slowly within 8 in. of floor (2 fists, 1 on top of other)			14	77.8	
6. *Trunk extension, backward*, lying on front, buttocks held down, with hands behind neck raise chin 18 in. from floor (forearm height)			13	72.2	Above average
Agility: 7. *Kneeling jump*, spring to feet, hold balance for 3 sec			12	66.4	Passing
8. *Jack spring*, touching hands to toes at least waist high, 5 times in succession			11	61.1	
9. *Agility 6-count exercise*, squat, extend legs backward, extend legs forward, flip over, return to squat-rest position, return to standing position (6 times in 20 sec)			10	55.5	Failing
Strength: 10. *Man lift*, pick up partner of own weight and place on shoulders for carry in 10 sec			9	50.0	Below average
11. *Stick body*, hold 30 sec, head on partner's knee, hands on hips			8	44.4	
12. *Extended press-ups*, from hands and toes without using elbows (women do forearm press-ups 20 sec)			7	38.8	Poor

Item	Check result*		Ring result		Class
	Pass	Fail	Passed	%	
Power:					
13. *Standing broad jump*, height plus 1 ft			6	33 3	Poor
Endurance					
14. *Floor push-ups*, 15 times (women do 30 from knees)			5	27 7	
15. *Straddle chinning*, 20 times (women do 10)			4	22 2	Very poor
16 *V-sit*, 60 sec			3	16 6	
17. *Breath holding*, 30 sec after running in place 120 sec, at 180 steps/min			2	11 0	
18 *Endurance hops*, in succession, 200 up and down, 200 straddle jump, 200 alternate stride, 50 on right toes, 50 on left toes, and as many full squat-jumps as possible			1	5.5	
Number of items passed =					

* Pass = v, fail = x

Fig 8-4 Cureton's Eighteen-item Adult Indoor Motor Efficiency Classification Test for Men and Women (without Apparatus)

high school girls was developed in two forms[13] One form of six items can be given on a pass or fail basis to girls in one class period, the other form includes eleven items plus the Harvard Step Test and requires two class periods to complete

OREGON MOTOR FITNESS TEST

A cooperative effort in Oregon involving a state committee, Oregon State College, and the University of Oregon developed the Oregon Motor Fitness Test.[14] The tests were constructed to measure arm and shoulder girdle strength and endurance, abdominal strength and endurance, muscular power, running speed and endurance, agility, and trunk flexibility.

The same test items can be used for boys at the elementary, junior high, and high school levels The test items for girls can be used at all three levels.

[13] Mary E. O'Connor and Thomas K Cureton, "Motor Fitness Tests for High School Girls," *The Research Quarterly*, 16 302–314, December, 1945

[14] Oregon State Department of Education, "Oregon Motor Fitness Test Battery," Salem, Ore., Oregon State Department of Education, revised, 1958

Boys:

1. Pull-ups: performed on a horizontal bar with palms outward. A score of one is counted each time the subject goes from a complete hang to a point where his chin is above the bar. Half counts are permitted but not more than four.

2. Jump and reach: standard vertical jump procedure (see Chapter 11).

3. Potato race, 160 yards: performed between three circles 1 foot in diameter. Circle 1 is drawn on the floor just behind the starting line, circle 2 is drawn on the floor 50 feet directly ahead of circle 1, and circle 3 is drawn beyond circle 2 exactly 70 feet from the starting line.

a. The subject runs to circle 2, picks up an eraser or 2- by 4-inch block of wood, returns to the starting line, and places it in circle 1.

b. The subject runs to circle 3, picks up block, and carries it to circle 1.

c. The subject picks up block in circle 1 and carries it back to circle 2.

d. The subject returns to circle 1, picks up the second block, and carries it back to circle 3.

e. The subject runs back to starting line.

The score is the elapsed time in seconds.

Girls:

1. Arm flexed hang: standing on a stool, the subject places her hands, palms outward and shoulder width apart, on a horizontal bar. Elbows are flexed to permit the chin to be held level with the bar. When the support is removed, the girl holds the chinning position as long as possible. Legs are held straight throughout. The score is the elapsed time in seconds that the girl is able to maintain *some* flexion in the elbow joint. When the elbow finally straightens, the watch is stopped.

2. Standing broad jump: performed in the standard manner. Measurements made to the nearest inch. The best of three jumps is recorded.

3. Crossed-arm curl-ups: The girl assumes a hook-lying position on the back with knees bent at a right angle and feet hip width apart, flat on the floor. The feet are held down by a partner. In the test the girl raises her trunk to a sitting position and returns to the back-lying position as many times as possible. The elbows must not help the body sit up; the feet must stay on the floor; resting between set-ups or bouncing up are not permitted. The score is the number of proper sit-ups.

Raw scores for the tests are related to T-score charts and norm tables for all three levels of boys and girls. Norm tables for elementary boys, elementary girls, secondary boys, and secondary girls appear in Tables A-19 and A-20, Appendix A. These tests have much in their favor. They are interesting to perform and require very little special equipment.

INDIANA MOTOR FITNESS TESTS

The Indiana Motor Fitness Tests were validated early in World War II by Karl Bookwalter against a twelve-item standard involving two or more

measures each of strength, velocity, motor ability, and endurance. They were originally designed for high school and college age men.[15] Subsequently tests were developed for high school girls and elementary school age children.

College Men and High School Boys. Four motor fitness indices are available for college men and high school boys. They vary in validity from .859 to .812 and may be selected according to local facilities and interest.

Motor fitness index I: (chins + push-ups) × vertical jump
Motor fitness index II: (chins + push-ups) × standing broad jump
Motor fitness index III: (straddle chins + push-ups) × vertical jump
Motor fitness index IV: (straddle chins + push-ups) × standing broad jump

Instructions for these tests are as follows (1) Chinning or pull-ups are accomplished in the usual manner, except that either grip may be used. (2) Push-ups are done from a prone position on the floor The subject must keep the back straight with only the toes and hands on the floor He pushes to a straight-arm position This counts one The subject then lowers the body until the chest touches the floor He continues the exercise, doing it as many times as is possible (Girls do the push-ups with the knees on the floor instead of the toes) (3) The vertical jump is the jump and reach test, better known as the Sargent Vertical Jump This is performed in the usual manner and scored as total inches jumped (see Chapter 11) (4) The standing broad jump can be performed at the edge of an outdoor pit or on a mat in the gymnasium Measurements are made to the nearest inch, and the best of three jumps is recorded (5) For the straddle chins, pairs of subjects of about the same height measure each other The person to be tested lies on his back gripping the fingers (sailor grip) of the partner who stands astride him. The bottom person, holding his body straight, chins himself as many times as possible

Raw scores are related to scale scores for any one motor fitness index.

Example

A college man scored as follows

Straddle chins	20 times	Scale score = 59
Push-ups	22 times	Scale score = 65
Standing broad jump	98 inches	Scale score = 75

Apply the formula for index IV

$$(59 + 65) \times 75 = 9300$$

Divide by 100 and round off to the nearest whole number The MFI score is 93.

This takes on meaning according to the rating scale, where six standard deviations of the distribution of scores are made.

[15] Karl W Bookwalter and Carolyn W. Bookwalter, *A Measure of Motor Fitness for College*, Bulletin of the School of Education, Indiana University, vol. 19, no. 2, March, 1943.

	Rating	Motor Fitness Index
A	Superior	85 and up
B	Good	59 to 84
C	Normal	33 to 58
D	Fair	7 to 32
F	Inferior	6 or less

Local norms can be established on a local distribution of scores by computing means and standard deviations for each test item The norm can be built according to standard deviation units removed from the mean.

High School Boys and Girls: In this variation of the Indiana Motor Fitness Test all boys and girls perform four items: straddle chins, squat-thrusts, push-ups, and vertical jump [16] The squat-thrusts are performed as a four-part exercise as rapidly as possible for a full twenty seconds From a standing position of attention the subject (1) drops to a squat-rest position with hands on the floor, (2) extends the feet to the rear keeping the body straight, (3) returns to the squat-rest position, and (4) stands up straight. This is done as a continuous exercise with a score of one given for each repetition of the exercise

Scoring is accomplished as follows:

$$\frac{(\text{Straddle chins} + \text{squat-thrusts} + \text{push-ups}) \times \text{vertical jump}}{10}$$

Thus a boy doing fourteen straddle chins, ten squat-thrusts, eighteen push-ups, and twenty-two inches in the vertical jump would have a raw test score of $(14 + 10 + 18) \times 22/10 = 924/10 = 92$ This raw score of 92 is then related to McCloy's Classification Index I, which considers age figured to the nearest year, height to the nearest inch, and weight to the nearest pound. All ages over seventeen years are calculated as seventeen years:

$$\text{McCloy's Classification Index} = 20\ (\text{age}) + 6\ (\text{height}) + \text{weight}$$

To continue the illustration, suppose that the boy with the raw score of 92 was 16 years old, was 65 inches tall, and weighed 128 pounds. The classification index would be $20(16) + 6(65) + 128 = 838$. Relating the classification index of 838 to the performance score of 92 on the norm table shows that the boy is rated "average" (see norm table, Table A-11, Appendix A) The same performance by a girl would have been rated "good" Two tables in the Appendix, however, would have to be consulted to arrive at this rating: first Table A-11 to figure the height-weight classification division; then Table A-12, Achievement Scales for Girls, to determine the specific rating.

Elementary School Boys and Girls: Working with the four items in the secondary school test, Franklin and Lehsten adapted it for elementary school

[16] State of Indiana, *Physical Fitness Manual for High School Boys,* Department of Public Instruction, Bulletin no. 136, 1944, p. 21.

children in grades four to eight[17] There is no change in test administration or scoring. Also, pupils are classified according to McCloy's Classification Index I The norm tables are different and may be found in Appendix A (Table A-13 for boys and Table A-14 for girls).

These tests have proved valuable over a number of years for screening elementary school age children They point up weaknesses in both physical capacity and general motor coordination Their use for pupil classification is justified

MINNESOTA PHYSICAL EFFICIENCY TEST

The Minnesota Physical Efficiency Test was designed for pupils in the secondary schools of Minnesota[18] It is composed of six selected items: standing broad jump, pull-ups, push-ups, sit-ups, vertical jump, and the Burpee test The test is given in two parts Group I consists of sit-ups, push-ups and the Burpee test Group II consists of the broad jump, vertical jump, and pull-ups Minimum essential norm tables with achievement levels are available for boys and girls grades seven to twelve

WASHINGTON ELEMENTARY SCHOOL PHYSICAL FITNESS TEST

Under the direction of Kirchner the Washington Association for Health, Physical Education and Recreation developed a test that can be given by the elementary classroom teacher[19] It offers the elementary school teacher an objective instrument with which to measure physical status, and it provides a means of comparing the individual child to the group The test is designed for follow-through action so that the parents will be more aware of the physical needs of their children

There are five tests in the battery standing broad jump bench push-ups, curl-ups squat-jumps, and a 30-yard dash The tests are completely described in the manual Moreover there are elaborate separate norm tables for boys and girls for each age six to twelve Pupils are classified as being superior, average, poor or very poor in each test item as well as in total physical fitness as measured by the sum of the five tests Emphasis is placed on the interpretation of test scores and the use of the Physical Fitness Report Card

The tests are carefully explained with clear drawings so the classroom teacher will have a minimum of difficulty in measuring power, strength and endurance, and speed

[17] C C Franklin and N G Lehsten, "Indiana Physical Fitness Tests for the Elementary Level (Grades 4 to 8) " *The Physical Educator,* 5 38–45, May, 1948

[18] *The Minnesota Physical Efficiency Test, A Guide for Instruction in Physical Education,* Curriculum Bulletin no 11, State of Minnesota Department of Education, St Paul, Minn , 1950, pp 42 48

[19] Glenn Kirchner, *Elementary School Physical Fitness Test: A Manual for Classroom Teachers,* 1959, copies available from author at Eastern Washington College, Cheney, Wash.

JCR TEST

This is a three-item test consisting of the vertical jump (J), chin-ups with the forward grip (C), and the 100-yard shuttle run (R), in which the subject runs a 10-yard course ten times, aided by backboards that help in making the 180-degree turns.[20] The test is designed to appraise total ability through skills involving the basic elements of power, strength, speed, agility, and endurance. Reliability is high, and validity is satisfactory. Scoring tables based on the 6-sigma scale are available from the basic reference.

NSWA TEST

Under the direction of its Research Committee, the National Section on Women's Athletics empirically set up eight test items designed to appraise "muscular control and coordination, speed, agility of movement, and strength to move the body . . ."[21] Norm scales were developed using three standard deviations below and above the mean. This was done after administering the test to over twenty thousand high school girls. The scaled scores for the battery are reproduced in Table A-18, Appendix A.

In the NSWA Test the physical performance items consist of the following[22]:

1. *Basketball throw for distance:* The ball is thrown from behind the restraining line and distance measured to the nearest foot. The best of two trials is recorded.

2. *Potato race:* The subject runs between two lines 30 feet apart. She runs first from line 1 to line 2 to pick up a 2- by 2- by 4-inch wooden block and then back to line 1 to place the block behind the line. She runs once more to line 2 to pick up the second block and returns to the start, crossing line 1 with the block in her hand. The best of two trials is recorded in seconds and fifths of a second.

3. *Sit-ups:* These are performed in the typical manner and scored according to the number executed.

4. *Push-ups:* These are done from the floor from the knee position and scored according to the number executed.

5. *Pull-ups:* These are done on a 3½-foot high horizontal bar from a starting position in which the feet are on the floor, knees bent at a right angle, and arms fully extended. Scored according to the number of pull-ups executed where the chest is brought to the bar.

6. *Squat-thrusts, ten seconds:* These are performed as rapidly as possible to give a measure of agility and scored according to the number of repeti-

[20] B. E. Phillips, "The JCR Test," The Research Quarterly, 18:12–29, March, 1947.

[21] Eleanor Metheny, "Physical Performance Levels for High School Girls," *Journal of Health and Physical Education*, 16:32–35, June, 1945.

[22] More specific details are given in the original source. See footnote 21.

tions performed in ten seconds (see description of squat-thrusts in Navy Standard Physical Fitness Test).

7. *Squat-thrusts, thirty seconds:* A partial measure of endurance, these are performed for thirty seconds and scored according to the number executed.

8. *Standing broad jump:* Performed from the floor, mat, or jumping pit in the typical manner and scored in feet and inches.

In order to meet local situations a short form of the test was devised. It is composed of five items: standing broad jump, basketball throw, potato race *or* ten-second squat-jump, sit-ups, and push-ups *or* pull-ups.

The NSWA Test affords the subject a well-rounded workout and appears to have considerable merit when used to estimate the motor fitness of high school girls. The scaled scores for each test item are totaled These are averaged for the eight tests given If, for example, the average of the eight tests was 50, then the girl would be about average because a scaled score mean of 50 is average

A.A.U. JUNIOR PHYSICAL FITNESS TESTS

For a number of years the Amateur Athletic Union of the United States has been promoting the physical fitness of boys and girls aged six to

BOYS

Event	Age group, years				
	6 7	8 9	10 11	12–13	14–15
Required events					
1. Sprints	40 yd, 9 sec	40 yd, 8 sec	50 yd, 8 sec	60 yd, 9 sec	100 yd, 14 sec
2 Walk and run	¼ mile, 5 min	½ mile 8 min	¾ mile, 10 min	1 mile, 11 min	1½ mile, 18 min
3. Sit-ups	8 times	12 times	16 times	20 times	25 times
4. Pull-ups	(modified) 3 times	(modified) 7 times	regular 3 times	regular 5 times	regular 6 times
5. Standing broad jump	3 ft	4 ft	5 ft	5 ft 6 in.	6 ft
Choose any one of these events:					
6. Push-ups	(modified) 5 times	(modified) 8 times	(modified) 13 times	regular 10 times	regular 12 times
7. Playground ball throw for distance	35 ft	65 ft	85 ft	100 ft	125 ft
8. Continuous hike for distance	1 mile	2 miles	3 miles	4 miles	5 miles
9. Running high jump	1 ft 6 in.	2 ft 3 in.	2 ft 9 in.	3 ft 3 in.	3 ft 9 in.

Fig. 8-5. A.A.U. Junior Physical Fitness Test Standards (Boys).

fifteen.[23] Community organizations, such as the YMCA, and a number of elementary schools have used these tests to create an awareness of fitness. They have been advanced by the Junior Chamber of Commerce and other school-community groups.

In these tests a boy or girl can win a certificate of merit from the A.A.U. by successfully completing six of nine events, as shown in Figures 8-5 and 8-6.

GIRLS

Events	Age group, years				
	6–7	8–9	10–11	12–13	14–15
Required events:					
1. Sprints	40 yd, 9 sec	40 yd, 8 sec	50 yd, 9 sec	60 yd, 10 sec	100 yd, 17 sec
2. Walk and run	¼ mile, 5 min	½ mile, 8 min	¾ mile, 11 min	1 mile, 13 min	1½ mile, 19 min
3. Sit-ups	8	12	14	16	18
4. Pull-ups	(modified) 3 times	(modified) 7 times	(modified) 8 times	(modified) 9 times	(modified) 10 times
5. Standing broad jump	3 ft	4 ft	4 ft 6 in.	5 ft	5 ft 6 in.
Choose any one of these events:					
6. Push-ups	(modified) 4	(modified) 7	(modified) 9	(modified) 10	(modified) 12
7. Playground ball throw for distance	20 ft	30 ft	40 ft	50 ft	75 ft
8. Continuous hike for distance	1 mile	2 miles	3 miles	4 miles	5 miles
9. Running high jump	1 ft 6 in.	2 ft 3 in.	2 ft 6 in.	2 ft 9 in.	3 ft

Fig. 8-6. A.A.U. Junior Physical Fitness Test Standards (Girls).

The chief advantage of the A.A.U. tests is their adaptability to any school situation; also they can be used with little motivation as self-testing activities. Moreover, any classroom teacher can understand them and use them to screen pupils low in physical capacity.

MILITARY TESTS OF PHYSICAL FITNESS

With the advent of World War II there was an immediate demand for men and women of a high level of physical capacity to engage in strenuous and fatiguing activity. All branches of the Armed Forces had some variety of a motor performance test designed to supplement the physical examina-

[23] The Amateur Athletic Union, 233 Broadway, New York.

tion and discover who among the military personnel lacked the vigor and endurance to fight from a foxhole, carry a pack, pilot an aircraft, etc The tests had wide use Some of them have been modified and are used today in the Armed Forces as well as by a number of college and university departments of physical education

U.S. Air Force Physical Fitness Test

The three-item test, for men aged eighteen to thirty-eight, was developed from some fifteen test items known to measure fitness elements [24] Alternate tests are available, one for indoor testing and the other for outdoor testing

Indoor test	*Outdoor test*
Sit-ups	Sit-ups
Pull-ups	Pull-ups
250-yard shuttle run	300-yard shuttle run

In these tests the sit ups are done in the typical manner, with a partner holding the ankles down The subject performs as many sit-ups as possible but no more than 112 Pull ups are done on the high horizontal bar with palms facing away from the body The 300-yard shuttle run is done on a smooth, level surface in a marked-off running lane 6 feet wide and 60 yards long A stake extending at least 2 feet above the ground is placed at both ends of the lane On the signal from the leader the subject runs toward the stake at the other end of the lane and runs back and forth for five trips (total 300 yards) Time is recorded in seconds In the 250-yard shuttle run (indoors) the subject runs over a course 6 feet wide and 25 yards long with 6-inch high turning blocks at the ends of the lane He runs back and forth for ten trips (total, 250 yards)

The number of sit-ups the number of pull-ups, and the running time in seconds are the raw scores These are changed to standard scores by using the scoring tables Once the standard scores have been found, they are totaled to provide a composite standard score From the same scoring table the composite standard score can be changed to a physical fitness rating (PFR) The PFR is classified as excellent very good, good, poor, or very poor

The example recorded in Figure 8-7 is for subject who did 63 sit-ups and 10 pull-ups and ran the shuttle-run in 49 seconds His standard scores ("Performance Scores') were 72, 58, and 68, respectively, yielding a total of 198 points, the composite score A score of 198 yields a PFR of 66, which gives the subject a rating of very good

The test is excellent for use with senior high school boys and college men, chiefly because it consists of only three easy to administer valid, and

[24] Air Force Manual 160-26 *Physical Conditioning*, Department of the Air Force, Washington, April, 1956, see pp 155–160

Computation of Performance Scores PFR Scores and Rating

Number of sit ups (Test Performances)	Number of pull ups (chinning) (Test Performances)	Shuttle Runs: Outdoor time in seconds	Shuttle Runs: Indoor time in seconds	Performance and PFR scale (Scoring Scales)	Composite performance scale (Scoring Scales)	Rating scale
114	24	41	41	100	300	EXCELLENT (E)
111				99	297	
108	23			98	294	
105				97	291	
102	22	42		96	288	
99	21			95	285	
96		43	42	94	282	
93	20			92	276	
90	19			90	270	
87				88	264	
84	18	44	43	86	258	
81				84	252	
78	17			82	246	
75				80	240	
73	16	45	44	78	234	
72				77	231	VERY GOOD (VG)
69	15	46	45	76	228	
63				74	222	
66	14	47		72	216	
60		48	46	70	210	
57	13	49	47	68	204	
54	12	50	48	66	198	
52				64	192	
51	11	51	49	62	186	GOOD (G)
48		52		60	180	
45	10			58	174	
		53	50	56	168	
42				55	165	
39	9			54	162	
36		54	51	52	156	
33	8			50	150	
31		55	52	48	144	
30	7	56	53	46	138	POOR (V)
		57	54	44	132	
27	6	58	55	42	126	
		59		40	120	
24	5	60	56	38	144	
		61	57	36	108	
21	4	62	58	34	102	
20	3	63	59	32	96	VERY POOR (VP)
18		64	60	30	90	
15		65	61	28	84	
12	2	66	62	26	78	
9		67	63	22	66	
6	1	68	64	18	54	
3		69	65	14	42	
1		70	66	10	30	

TEST RECORD

TEST NUMBER	EXAMPLE Test Perf	EXAMPLE Perf Score	1 Test Perf	1 Perf Score	2 Test Perf	2 Perf Score	3 Test Perf	3 Perf Score	4 Test Perf	4 Perf Score	5 Test Perf	5 Perf Score	6 Test Perf	6 Perf Score	7 Test Perf	7 Perf Score	8 Test Perf	8 Perf Score	9 Test Perf	9 Perf Score	10 Test Perf	10 Perf Score
Sit Ups	63	72																				
Pull Ups (Chinning)	10	58																				
Shuttle Run (Outdoor or indoor)	49	68																				
Composite Perf Score		198																				
PFR Score	66																					
Rating	VG																					

Fig 8 - USAF Physical Fitness Test Scoring Scales and Record Form

reliable test items.[25] Furthermore, it can be taken in a short period of time. The Strategic Air Command (SAC) has used it extensively as the chief element on the Individual Progress Record (Physical Conditioning); the men like to keep their scores; and the PFR means something specific to the air base commanders.

Marine Corps Physical Readiness Test

This test for men up to age forty was designed to determine whether a marine meets a minimum standard of physical readiness.[26] It consists of six tests, which are scored as pass or fail.

Test Item	*Satisfactory Score*
Chin-ups from bar	3
Push ups from floor	21
Two-minute sit-ups	25 in two minutes
One-minute squat-thrusts	15 in one minute
Standing broad jump	6 feet
Half-mile run or jog (no time limit)	Ability to run or jog a half mile without stopping

It should be pointed out that these requirement are minimum standards, and will therefore be easy for many college men to meet.

Navy Standard Physical Fitness Test

The United States Navy developed a test to appraise the physical fitness of Navy personnel.[27] Its use, however, is broader than this The individual needs of the men, their motivation to keep fit, their progress, and how they compare with other Navy men in other service activities are major physical fitness program objectives.

There are five items in the test battery. They are described briefly as follows:

1. *Squat-thrusts:* From a squat position with the hands placed on the floor, the subject thrusts the legs backward in an extended manner so the body is in a push-up position From here the subject returns his legs to the squat position He then stands up straight with chest and head up. The subject is scored on the number of complete squat-thrusts completed in one minute.

2. *Sit-ups:* These are performed in the typical manner, continuously until the subject is exhausted.

[25] Leonard A Larson, "Some Findings Resulting from the Army Air Forces Physical Training Program," *The Research Quarterly,* 17:136-144, May, 1956.

[26] Marine Corps Manual, chap 24 and Marine Corps Order 6100 3A, *Physical Readiness Test,* Apr. 9, 1958, Department of the Navy, Headquarters United States Marine Corps, Washington 25, D C

[27] *Physical Fitness Manual,* U.S Navy, Bureau of Naval Personnel, Training Division, Physical Section, 1943, chap. 4.

3. *Push-ups:* These are done from a prone position with the chest on the floor A point is scored each time the subject completes a push-up Points are scored only when the body is held straight

4 *Squat-jumps* In this exercise the subject stands with the feet about 6 inches apart and the heel of the left foot even with the toes of the right foot The hands are placed on top of the head palms down and fingers interlaced The subject performs the following movement as many times as he can until exhausted

a Drop to a squat on the right heel

b Immediately spring upward with *both knees straight* While in the air, change the position of the feet so the right foot is in front then drop to a squat position over the left heel

c Repeat movement continuously interchanging the feet each time while off the floor (This event should be practiced prior to testing until it can be done with ease)

5 *Pull-ups* These are performed on the horizontal bar in the typical manner that is, the chin must be raised above the bar, the arms must be fully extended, and no kipping movement or resting is permitted

The Navy test has had wide use both in the service and in higher educational institutions Each test item can be related to a T score table which was standardized on well-conditioned Navy personnel (see Table A-17,* Appendix A)

Army Physical Efficiency Test

The U S Army test was validated on the physical performance of conditioned and nonconditioned military personnel [18] The original test consisted of ten items designed to measure the essential elements of physical fitness Because of the time involved in administering the test it was later reduced to five items: (1) pull-ups (2) squat-jumps, (3) push-ups (4) sit-ups, and (5) a 300-yard shuttle run Should it be necessary to give the test indoors, a 250-yard shuttle run or the sixty-second squat-thrust can be substituted for the 300-yard run The test items are given in the same fashion as in the Navy and Air Force tests Scoring tables are available in the original field manual source

Military Academy Tests

Each of the service academies has its own physical proficiency test These items have been proved valid by their use in the various test batteries of the Armed Forces The *United States Military Academy Physical Efficiency Test* is used for admission and retention at West Point [19] The test battery,

[18] Field Manual FM 21-20, *Physical Training*, Department of the Army, Washington, January, 1946

[19] Department of Physical Education, United States Military Academy, West Point, N Y.

available from the academy, varies from year to year and includes such standard items as sit-ups, softball throw for distance, push-ups, and 100-yard shuttle run The *Air Force Academy Physical Aptitude Examination* includes a series of physical aptitude exercises designed to measure strength, coordination, endurance, speed, and agility[30] It is used to screen candidates for admission to the academy Later, the physical fitness test comprises a portion of the term grade in physical education courses

WAC Physical Fitness Test

This test was designed to evaluate the effect of the Womens Army Corps physical fitness program during World War II[31] It consists of (1) floor

Raw score	Points scored	Classification	
20	100	90-100	Excellent
19	98	80-89	Good
18	96	65-79	Average
17	94	45-64	Poor
16	92	0-44	Very poor
15	90		
14	88		
13	86		
12	84		
11	82		
10	80		
9	75		
8	70		
7	65		
6	60		
5	50		
4	40		
3	30		
2	20		
1	10		
0	0		

Fig 8-8 WAC Physical Fitness Rating Scores and Classification

push-ups or push-ups from the knees if the woman is incapable of doing full dips (2) sit-ups with feet on floor and knees at a right angle (3) wing-lifts (done from a prone position with feet pointed and hands interlaced behind the neck and body arched so chest and legs raise off the floor, subject does as many as possible in one minute), and (4) squat-thrusts, as many as possible in thirty seconds The test is useful for college women, but norms need to be developed for high school girls Norms suitable for college women are found in Figure 8-8

[30] Department of Physical Education United States Air Force Academy Colorado Springs, Colo

[31] *WAC Physical Fitness Test*, War Department, Washington, 1943 Available from U S Government Printing Office

A girl who does fifteen push-ups would have a point score of 90, which is considered excellent Likewise a raw score of fifteen squat-thrusts in thirty seconds would yield the same rating A composite score and total classification is determined by averaging the raw scores

Appraisal of Motor Fitness Tests

No test is perfect, and few tests are completely satisfactory One should not expect too much from most measurements At best they are screening devices and guides for future action Their value lies in their proper use as a means of improving programs There is a right measure for the right place and a wide variety of satisfactory tests (see Figure 8-9)

In reviewing motor fitness tests and their integral parts it soon becomes obvious that a number of measures employ similar test items designed to measure some phase of physical capacity For example such items as pull-ups, push-ups, and sit-ups appear often as a means of appraising strength and endurance This tends to give importance to the items themselves because most tests were validated with them

There is a relatively high correlation between test batteries Brown discovered, after carefully examining twelve batteries of tests and establishing a criterion score, that college men obtained scores on the criterion that correlated above 80 with Rogers Strength Index McCloy's Motor Ability Test, Larson's Motor Ability Test (Outdoor) and the Indiana Motor Fitness Test [32]

A number of significant correlations between the various Armed Forces fitness tests have been worked out and reported by Matthews, Shay, and Clarke [33] Here are some of the findings

1 Squat-thrusts (number in one minute 79 with Navy test and 62 with Army test

2 Shuttle-run (300 yards scored in seconds) 81 with Army test, 74 with Air Force test -72 with Strength Index -66 with Physical Fitness Index and -61 with Navy test

3 Multiple correlations

a U S Air Force Physical Fitness Test 82 300-yard shuttle run and sit-ups

b Navy Standard Physical Fitness Test 93 squat-thrusts, squat-jumps, and 300-yard shuttle run

c Army Physical Efficiency Test 95 300-yard shuttle run, sit-ups, and squat-jumps

d Strength Index, 98 leg lift and squat-thrusts

e Physical Fitness Index, 99 leg lift squat-thrusts and sit-ups

[32] Howard S Brown, "A Comparative Study of Motor Fitness Tests" *The Research Quarterly*, 25 8 12, March, 1954

[33] Donald K Matthews Clayton I Shay and H Harrison Clarke, "Relationship between Strength Loss in Pack Carrying and Certain Motor Physical Fitness Criteria," *The Research Quarterly*, 26 426 430, December, 1955

Test of physical fitness	Elementary school	Secondary school	College men	College women	Male	Female
Muscular strength and endurance tests:						
Cable-tension Strength Tests		x	x	x	x	x
Drop-off Index		x	x	x	x	x
Kraus-Weber Test	x	x	x	x	x	x
Larson Muscular Strength Test			x		x	
McCloy's Athletic Strength Test	x	x			x	
Physical Fitness Index (original)	x	x	x	x	x	x
Physical Fitness Index (Oregon Simplification)	x	x			x	x
Roger's Strength Index	x	x	x	x	x	x
Motor test batteries:						
AAHPER Youth Fitness Test	x	x			x	x
A.A.U. Junior Physical Fitness Test	x	x			x	x
Air Force Academy Physical Aptitude Examination		x	x		x	
Army Physical Efficiency Test			x		x	
California Physical Performance Test	x	x			x	x
Indiana Motor Fitness Test (Elementary)	x				x	x
Indiana Motor Fitness Test (High School)		x			x	x
Indiana Motor Fitness Test (Men and Boys)		x	x		x	
JCR Test		x	x		x	
Marine Corps Physical Readiness Test		x	x		x	
Minnesota Physical Efficiency Test	x	x			x	x
Motor Fitness Screen Test for High School Girls		x				x
National Section on Women's Athletics Test		x		x		x
Navy Standard Physical Fitness Test			x		x	
New York State Physical Fitness Test	x	x			x	x
Oregon Motor Fitness Test (Boys)	x	x			x	
Oregon Motor Fitness Test (Girls)	x	x		x		x
University of Florida Physical Fitness Test			x		x	
University of Illinois Motor Fitness Test			x		x	
University of Maryland Motor Fitness Test			x		x	
U.S. Air Force Physical Fitness Test			x		x	
U.S. Military Academy Physical Efficiency Test		x	x		x	
Washington Elementary School Physical Fitness Test	x				x	x
WAC Physical Fitness Test		x		x		x
Yale University Physical Fitness Test			x		x	

Fig. 8-9. Classification of Physical-Motor Fitness Tests According to Suitability.

Certainly a wide choice of motor performance measures that can be used to assess physical fitness is available to the physical education teacher. There is little excuse for failure to use *some* one of these means to evaluate the program. Frequently, one hears the old argument that "tests are too difficult to give," "they take up too much of my class time," "they are inadequate for girls," or "there is little for the elementary school level." Such criticism is simply unfounded. A review of the research supports the opposite point of view. Evidence abounds that proper testing is part of the program and should require part of the class time. Tests are available at all levels and can be administered by almost any interested school teacher. Moreover, many measures are suitable for girls and women. One has only to experiment with them and, if necessary, make modifications at the local level. Obviously, this is a most neglected area in many schools. Yet girls exposed to an appraisal program enjoy it and profit from it every bit as much as boys.

Measures Suitable for Elementary School Use

An examination of Figure 8-9 reveals that there are a number of suitable measures for use with elementary school boys and girls. Tests that have been used over the years and are of unquestionable value include the Physical Fitness Index, the Kraus-Weber Tests, the Oregon Motor Fitness Test, the Indiana Motor Fitness Test, and the AAU Junior Physical Fitness Test. The newer measures, such as the AAHPER Youth Fitness Test and the state tests from New York, California, and Washington, are finding wide usage.

The best place to screen pupils for low fitness is in the primary and intermediate grades. Remedial efforts often prove especially fruitful with children in the early years when parents are more than usually interested in cooperating with school personnel. In New York State it is stressed that the value of the pupil's score card from the New York State Physical Fitness Test be made known to parents. The same approach is taken in using fully the AAHPER Personal Fitness Record.

Research indicates that more evaluation of children's progress in physical education is needed. This is especially true if program values are to be brought before the school administration and the public. It is encouraging, however, to find that scientific evaluation practices are being engaged in nationally. In a study of physical education in urban elementary schools carried on by the U.S. Office of Education, it was found that out of a sample of 523 city school systems, 25 per cent (129) indicated that children are given tests for physical fitness.[31] Although the items included in the tests vary from community to community, certain aspects of physical fitness appear to be of particular concern in many of the school systems reporting: agility,

[31] U.S. Department of Health, Education, and Welfare, Office of Education, *Physical Education in Urban Elementary Schools*, Bulletin 1959, no. 15, 1959, pp. 38-41.

balance, endurance, flexibility, speed, and strength. Among the tests frequently reported were the Kraus-Weber Tests, A.A.U. test, Iowa Revision of the Brace Motor Ability Test, and the Minnesota Physical Efficiency Test.[35]

SELECTED REFERENCES

Brown, Howard S.: "A Comparative Study of Motor Fitness Tests," *The Research Quarterly*, 25:8 19, March, 1954.

Bucher, Charles C., and Dominick A. Taddonio: "The Relationship between the Physical Fitness Ratings of Aviation Cadets and Certain Early Life Experiences Pertaining to Physical Activity," *The Research Quarterly*, 30:136-143, May, 1959.

Clarke, H Harrison: "The Physical Educator's Role in Physical Fitness," *Northwest Medicine*, 10:56 60, January, 1960.

Henry, Franklin M.: "Influence of Measurement Error and Intra-individual Variation in the Reliability of Muscle Strength and Vertical Jump Tests," *The Research Quarterly*, 30:155 159, May, 1959.

Hunsicker, Paul: "Physical Fitness Tests," *Journal of Health, Physical Education and Recreation*, 28:21 23, September, 1957.

—— : "AAHPER Youth Fitness Project," *Journal of Health, Physical Education and Recreation*, 28:17, November, 1957.

—— : "AAHPER Physical Fitness Test Battery," *Journal of Health, Physical Education and Recreation*, 29 24 25, September, 1958.

Rarick, Lawrence, Katherine Gross, and Mona J. Mohns: "Comparison of Two Methods of Measuring Strength of Selected Muscle Groups in Children," *The Research Quarterly*, 26:71 79, March, 1955.

Scott, M. Gladys, and Marjorie Wilson: "Physical Efficiency Tests for College Women," *The Research Quarterly*, 19:62 69, May, 1948.

[35] *Ibid.*, p. 39.

CHAPTER 9

Measures of Posture and Body Mechanics

Posture the perpendicular line connecting heaven and earth

Martha Graham

The study of posture involves the position of the various body segments at any given moment It considers the mechanics of movement, especially the more fundamental movements of walking running and sitting, as well as those related to daily tasks such as bending, stooping pushing, and lifting There is even an optimum sleeping posture Individual body build and its influence on behavior together with anthropometric measurements in general, are related to posture Studies in kinesiology and physiology therefore, are basic to the subject of posture Structure precedes function, and the function of the whole organism is related to the sum effect of its separate parts

Posture and Health

There is a long history of effort to measure and relate posture to health and well-being This has been viewed chiefly as a physiological study, with some attention to mental and emotional health Research indicates more than a theoretical relationship between health and posture Logically one might be expected to perform below par when internal organs (the viscera) nerves, blood vessels etc , are displaced because of the position of the body Yet the human organism with its remarkable powers of adaptation to untold stresses and strains, in many cases shows no outward signs of reduced efficiency even when the static posture or body mechanics is below standard This can mean two things Either the standards are less than completely scientific, or too much value is attached to posture It is not uncommon to find educators who would do away with posture training on the basis that many postural defects are temporary and represent a feeling or attitude of the moment The evidence, however, is quite the other way Postural defects, when uncorrected, tend to remain or progressively worsen over the years [1] As Howland has pointed out postural habits established in the early years carry over to adulthood, with the corresponding difficulties of align-

[1] This has been brought out repeatedly by Drs Kendall and Kendall at the Children s Hospital, Baltimore, Md

ment and movement.[2] The early orthopedic work of Kraus and Weber supports this statement.[3]

Certainly there is a substantial body of evidence in support of proper posture. Cureton reviewed much of this evidence and showed a relationship between posture and health status.[4] He warned that posture may have esthetic and social values "and in extreme cases it undoubtedly is a barometer of physical fitness, but it cannot be depended upon as an index of fitness."[5]

Moriarity and Irwin studied school children in Massachusetts schools and found a significant association between poor posture and numerous physical, mental, and emotional disorders.[6] Children with habitual poor posture had more disease, fatigue, underweight, self-consciousness, fidgeting, hearing defects, restlessness, timidity, and asthma. Fox, although pointing out that the causes of dysmenorrhea are numerous, states that a lack of abdominal strength together with a sway-back condition may be a cause of this condition in college women.[7] Flint also found a definite relationship between chronic backache and the strength of the abdominal and back musculature.[8] Further work has been done in this area by Larson, at the Children's Hospital in Iowa City, where it was demonstrated in one study that 96 per cent of cases of low back pain were relieved after a course of progressive exercises. Perhaps the most publicized work evaluating posture and relating it to health status was carried on by Kraus and Hirschland at New York University Bellevue Medical Center, where 1,000 cases of low back pain were evaluated.[9] Trunk muscular imbalance and weaknesses were present in varying degrees in all cases. This finding is in keeping with the report of the Baruch Committee on Physical Medicine, which maintains that many ailments and disabilities, including severe and crippling pain, may be caused by abnormal posture.[10]

[2] Ivaclare Sprow Howland, *Body Alignment in Fundamental Motor Skills*, New York, Exposition Press, 1953.

[3] Hans Kraus and S. Weber, "Evaluation of Posture Based on Structural and Functional Measurements," *Physiotherapy Review*, 25:267–275, November–December, 1945.

[4] Thomas K. Cureton, "Bodily Posture as an Indicator of Fitness," *The Research Quarterly* (Supplement), 12:348–382, May, 1941.

[5] Thomas K. Cureton, *Physical Fitness Appraisal and Guidance*, St. Louis, The C. V. Mosby Company, 1947, p. 47.

[6] Mary J. Moriarity and Leslie W. Irwin, "A Study of the Relationship of Certain Physical and Emotional Factors to Habitual Poor Posture among School Children," *The Research Quarterly*, 23:215–221, May, 1952.

[7] Margaret G. Fox, "The Relationship of Abdominal Strength to Selected Postural Faults," *The Research Quarterly*, 22:111–144, May, 1951.

[8] M. Marilyn Flint, "Effect of Increasing Back and Abdominal Muscle Strength on Low Back Pain," *The Research Quarterly*, 29:160–171, May, 1958.

[9] Hans Kraus, "Diagnosis and Treatment of Low Back Pain," *General Practitioner*, 4:45–60, April, 1952.

[10] *Report of the Baruch Committee on Physical Medicine*, 597 Madison Ave., New York, Apr. 1, 1945.

Perhaps more than anything else, functional postural defects are concomitant symptoms of illness and poor health. The malnourished child, for example, who does not have the caloric intake to supply needed energy simply does not have the strength to hold himself up, poor posture is the inevitable result. A depressed mental attitude, chronic fatigue, or an intestinal parasite might produce the same result. Klein and Thomas studied grammar school children and showed that, when body mechanices improved, there was a corresponding improvement not only in health and physical efficiency but in schoolwork as well.[11] Many classroom teachers would readily verify this.

Posture Standards

Before one can appraise posture and relate it to health status and personal efficiency, it is necessary to have some standard or guide to refer to. In the first place, it is difficult to define any exact posture that will be suitable for all sizes and builds of men and women. No two persons are alike in the relationships of their anatomical parts. From a mechanical point of view, however, ideal posture is that in which the various segments of the body—head, neck, chest and abdomen—are all balanced vertically one on the other so the weight is borne mainly by the bony framework, with a minimum of effort and strain on the muscles and ligaments. This is when the long axis, seen in profile, is essentially a vertical line instead of a zigzag. This is straight-line posture, maintained at ease with a minimum of effort (see Figure 9-1). More specifically, in proper posture the straight line would fall approximately from the lobe of the ear through the center of the shoulder, through the hip and greater trochanter, just behind the knee cap and through the outside ankle bone to the floor. This centralization around a vertical axis is important to balance, for a departure from the vertical shifts the line to other parts of the body away from the head. This calls for a restatement of the vertical axis and a physical compensation in order to maintain it. This ability to refocus and compensate is the result of previous training in strength, flexibility, and sensitivity.

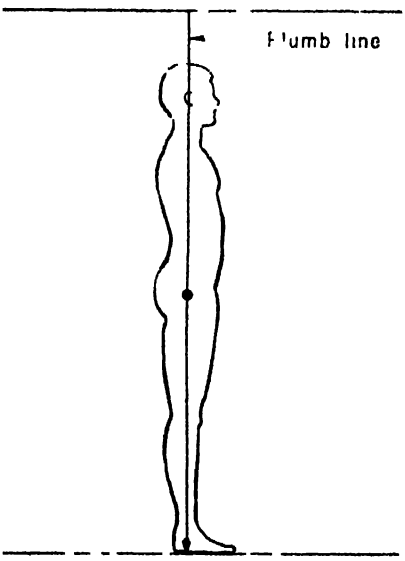


Fig 9-1 The "Straight Line" or Gravity Test

[11] Armin Klein and Leah C. Thomas, *Posture and Physical Fitness*, U S Department of Labor, Children's Bureau Washington, 1931

The straight or gravitational line was used as a standard in most early measures of posture. A person could be quickly measured by standing up to a plumb line dropped from the ceiling. The measure is subjective, but it still has considerable value. Opinions vary as to just where the line of gravity falls with most persons. Steindler feels that the gravitational line falls ahead of the ankle joint by about 4 centimeters.[12] Hellebrandt places the line of gravity 5.08 centimeters in front of the ankle joint.[13] Kendall, Kendall, and Boynton concur generally.[14] Fox and Young found that the gravitational line lay a mean distance of 0.95 centimeter in front of the anterior border of the tibia at the ankle.[15] There is good agreement that the mastoid bone or lobe of the ear is in line with the tip of the acromion process of the shoulder. Phelps, Kiphuth, and Goff, however, place little value on this point because they feel that the mobility of the shoulders is too great in most persons and would, therefore, distort any straight line.[16] Wells feels that more research, involving X-ray examination of straight and curved spines, is needed to obtain the hereditary pattern for any person.[17] After studying college women, she concluded that the spinal structure is not simple to classify. Her research indicates that there are two classes of spinal structure: (1) the humanoid type, with a long posterior concavity extending well up into the thoracic region, and (2) the anthropoid type, with a posterior convexity extending well down into the lumbar region.[18]

The Brunnstrom study supports the gravity-line approach to posture appraisal.[19] Lowman calls attention to the maintenance of proper body balance in relation to the vertical line.[20] Davies accepts this approach, not only in terms of posture measurement, but for teaching children to improve their posture patterns.[21] Rasch and Burke point out that posture standards

[12] Arthur Steindler, *Kinesiology of the Human Body under Normal and Pathological Conditions*, Springfield, Ill., Charles C Thomas, Publisher, 1955.

[13] Frances Hellebrandt *et al.*, "The Location of the Cardinal Anatomical Orientation Planes Passing through the Center of Weight in Young Adult Women." *American Journal of Physiology*, 121:465–470, 1938.

[14] Henry O. Kendall, Florence P. Kendall, and Dorothy A. Boynton, *Posture and Pain*, Baltimore, The Williams & Wilkins Company, 1952, p. 204.

[15] Margaret G. Fox and Olive G. Young, "Placement of the Gravitational Line in Anterio-posterior Standing Posture," *The Research Quarterly* 25:277–285, October, 1954.

[16] W. W. Phelps, R. J. H. Kiphuth, and C. W. Goff, *The Diagnosis and Treatment of Postural Defects*, 2d ed., Springfield, Ill., Charles C Thomas, Publisher, 1956, p. 172.

[17] Katherine F. Wells, "What We Don't Know about Posture," *Journal of Health, Physical Education and Recreation*, 29:31–32, May–June, 1958.

[18] Katherine F. Wells, "An Investigation of Certain Evolutionary Tendencies in the Female Human Structure," *The Research Quarterly*, 18:260–270, December, 1947.

[19] S. Brunnstrom, "Center of Gravity Line in Relation to Ankle Joint in Erect Standing," *Physical Therapy Review*, 34:109–115, March, 1954.

[20] Charles L. Lowman, "Faulty Posture in Relation to Performance," *Journal of Health, Physical Education and Recreation*, 29:14–15, January, 1958.

[21] Evelyn A. Davies, *The Elementary School Child and His Posture Patterns*, New York, Appleton-Century-Crofts, Inc., 1958.

need scrutinizing; that many scientific-sounding devices, such as conformateurs, comparagraphs, schematographs, and silhouetteographs have weaknesses; and that, although they do provide data about body alignment, "the *significance* of this information rests upon subjective and arbitrary assumptions implicit in the judgment of the measure."[22]

Although experimenters disagree on exact posture standards, most agree on some kind of erect, relatively straight-line body position for the standing person. Objective methods are encouraged. Even simple inspection measures can be made more objective by the use of check lists and rating scales. Pure observation has value when the observer is trained to observe carefully and knows what to look for. In short, the validity of posture appraisal methods can be improved so posture screening examinations will be more fruitful.

Posture Evaluation in the Elementary School

The chief purpose in appraising the posture of school children is to select pupils who deviate sufficiently from the standard to need help. The sooner this help arrives, the better. Thus the time to begin the program is in the primary grades, where all too frequently weak musculature and poor posture is the result of such underlying causes as malnutrition, fatigue, disease, underactivity, and personal attitudes of depression, withdrawal, and insecurity.

The role of the elementary classroom teacher is to appraise posture from day to day and to do what she can within the limits of her knowledge and understanding. Evaluation techniques, therefore, must be simple and possibly less objective than the more complicated standardized measures. The important thing, however, is that whatever the teacher detects or suspects is followed up by a pupil visit to the health service office, where the school physician can judge the problem and, if necessary, alert the physical education teacher, parent, and family physician. This is health education at its best.

In the elementary grades simple evaluation and instruction go hand in hand. Needed postural changes are frequently brought about when children hear, see, and *feel* what optimum posture is.[23] Both the mirror and the photograph give boys and girls a view of themselves. It has always been good practice to have them back up against a flat wall and feel with their hands how much space there is between the hollow of the back and the wall. Then, when they "stand tall" and walk away from the wall to a mirror they can reinforce how they feel with what they see. The same thing can be done with the plumb line dropped from the classroom ceiling, or a straight window pole:

[22] Philip J. Rasch and Roger K. Burke, *Kinesiology and Applied Anatomy,* Philadelphia, Lea & Febiger, 1959, p. 331.

[23] Carl E. Willgoose, *Health Education in the Elementary School,* Philadelphia, W. B. Saunders Company, 1959, p. 217.

1. The teacher drops a plumb line or holds a window pole or straight stick in a vertical position.
2. The pupil to be checked stands with his side adjacent to the pole.
3. Satisfactory posture is determined when the vertical line of the pole falls approximately through the lobe of the ear, the tip of the acromion process of the shoulder, the hip joint, just behind the kneecap, and the external ankle bone of the foot.

This Bancroft vertical test is weak in reliability, but it is one of the fastest screening techniques available for teachers who lack training in the more complicated devices. An experienced teacher, working with the school nurse or physical education instructor, is frequently able to judge standing posture and body alignment in motion well enough to detect the more serious faults. Where the postural elements are carefully defined, reliability improves. Cureton has shown that posture can be judged quite satisfactorily by school experts adequately trained in the principles of body mechanics.[24]

Most evaluation techniques work well in estimating the body mechanics of elementary children. Informal observations, self-check, pictures, cumulative records, anecdotal records, etc., are all useful.

Postural Analysis through Observation

Observation of school children may be formal or informal. Informal checks can be made about the school of various pupils as they work at their desks, eat their lunches, play on the playground, etc. Formal checks require more carefully planned observations. In such cases a testing procedure for individual pupils is usually established. One pupil at a time may be observed and rated for stance and gait.

The position of the head and neck, shoulders and chest, spine, pelvis, legs, ankles, and feet may be observed routinely in either a formal or informal manner. In many respects, the most productive observations are made by health and physical education personnel working with regularly scheduled pupils from day to day.

The head and neck should be held in an erect, well-balanced position so the eyes are level with the horizontal plane.

The shoulders and chest should meet the neck in an erect position with a slight adduction of the shoulders. There is a semielevated position of the ribs and sternum that takes up the slack in the abdominal muscles and provides a good base for their action.[25]

The spine, with its cervical concavity, thoracic convexity, lumbar concavity, and sacral convexity, when viewed laterally presents the picture of a smooth, flat S, which is not exaggerated in any area. The position of the hips, controlled primarily by the major abdominal muscles, determines in large part the extent of the lumbar concavity. This sway-back condition

[24] Thomas K. Cureton, "Bodily Posture as an Indicator of Fitness," p. 363.
[25] Philip J. Rasch and Roger K. Burke, *op. cit.*, p. 339.

is not difficult to observe and is often spotted by persons who work with children in schools, playgrounds, and beaches.

The *abdominal wall* should be straight or slightly convex. It should not sag. However, in the small child up to about seven years of age the abdomen will usually appear convex. As the child matures, this becomes reduced.

The *pelvis*, which has much to do with determining posture, can be observed informally by carrying out the following procedure:

As the subject is observed, draw an imaginary line from the sacrolumbar joint across the top of the symphysis pubes and extend the line down to the floor.

The angle formed by this line and the floor should approximate 55 to 60 degrees in the typical person.

The anterior superior spine of the ilium is *directly over* the symphysis pubis.

The *legs* should appear relatively straight with the kneecaps pointing directly ahead so there is no inward or outward rotation occurring at the hip or foot joints.

The *ankles and feet* need very careful observation both in the standing position and in walking and running. Essentially, straight ankles, as viewed from the rear, meet a foot that neither toes in nor toes out to an appreciable degree: that is, it is pointed ahead with the weight in walking transferred mainly from the heel, along the outside of the foot, to the forward arch.

Check Lists and Rating Scales. When the check list or rating scale is used, the observation becomes more accurate. Also, when posture and body mechanics is appraised by requiring the subject to walk about, a better picture of functional posture is presented.

Davies, in evaluating the posture of college women, used the check list in conjunction with the judgment of posture judges.[26] Divergencies included in the posture examination were as follows:

A. Gait (subject walking to and from judge):
Predominance of weight carried on one foot
Swings legs around instead of through
Feet everted
Ankles pronated
Weight on heels
Tibial torsion
Legs bowed
Knees knocked
Exaggerated hip swing
Exaggerated arm swing
Exaggerated bounce
Dragging or scuffing feet

B. Hip prominent

[26] Evelyn H. Davies, "Relationship between Selected Postural Divergencies and Motor Ability," *The Research Quarterly*, 28: 1–4, March, 1957.

C. *Trunk list*
D. *Shoulder rigid*
E. *Shoulder high*
F. *Head tilt*
G. *Scapulas prominent*
H. *Rotation*
I. *Restricted flexibility* (subject standing, bending forward from hips, returning slowly to erect position, then bending from side to side as far as possible)
J. *Lateral view* (judge standing at side while subject walks and then stands)
K. *Body lean*
L. *Shoulders forward*
M. *Thoracic curve:*
Decreased
Increased
N. *Lumbar curve:*
Decreased
Increased
O. *Hips forward*
P. *Abdomen protruding*
Q. *Knees hyperextended*

A check mark (✓) is placed after each divergency noted. They are then totaled for each subject.

Scott and French have illustrated a useful diagnostic body mechanics test developed at the State University of Iowa, which involves a three-point scale (3, good; 2, fair; 1, poor), to be used in rating subjects in such activities as running, sitting, stooping, reaching, stair climbing, carrying a load, pushing, and pulling.[27] In this test ten chairs are placed in a line, one in front of the other, with some space between. The subjects sit down. Names of subjects are recorded on the score sheet in the order in which they are sitting in the line of chairs. All subjects are dressed in leotards or swim suits and are barefoot. The rater stands to the side of the row of chairs. The following procedure is carried out:[28]

1. Each subject in turn walks a few steps toward the examiner, turns and walks away again. Foot alignment and pronation may be noted here.
2. Each subject in turn walks a few steps forward (with side to the examiner). Heel contact, weight transfer, and toe drive may be noted here.
3. Each subject stands in line while the examiner moves down the line rating standing posture.
4. Two or three subjects at a time walk back and forth beside the line of chairs while the examiner rates the walking posture. By having more than one walk at a time self-consciousness and unnatural gait are avoided.

[27] M. Gladys Scott and Esther French, *Measurement and Evaluation in Physical Education,* Dubuque, Iowa, Wm. C. Brown Publishing Co., 1959, pp. 227–229.

[28] *Ibid.,* pp. 227–228. Used by permission; mimeographed copies of the test and the three-point rating scale can be obtained from the Women's Department, Division of Physical Education, State University of Iowa, Iowa City, Iowa.

5. The subjects sit in the chair in a natural sitting position for rating. Each rises and sits again to be judged on balance and movements.

6. The subject walks up, then down, real stairs. It is suggested that these be wide stairs in order to accommodate two persons at a time.

Additional tests are set up in a similar manner. The secret of a realistic appraisal, according to Scott and French, is to provide properties or a setting in order to make all movements seem natural to the subject and examiner.

There are other ways to sharpen the observation technique. The observed posture can be given a specific rating or score and appraised according to a set key. The following rating sheet is an example. It is adapted from a chart used by Cureton and used here by permission.

Item	Specifications for rating	Scores, 0–10 points
I. Head posture	Note angulation of imaginary line from center of base of neck through tragus of the ear.	
II. Shoulder posture	Note position of shoulder with relation to the base of the neck and sternoclavicular notch.	
III. Chest posture	Note chest carriage in relation to body as a whole, the slope of the manubrium, subcostal angle, slope of ribs.	
IV. Abdominal posture	Note relative alignment of abdomen to chest, fist test to abdomen for control and development of muscles, amount of fat over muscles.	
V. Upper back	Note kyphotic curve; note after asking subject to flatten it; note while subject is bent over from the waist.	
VI. Lower back	Note lordosis curve; note after asking subject to flatten it; note while subject is bent over from the waist.	
VII. Pelvis	Note type, mark slope of iliac crest with flesh pencil; note tilt of pelvic basin as a whole.	
VIII. Knees	Note position of joint at rear of knee; note angulation of tibia to plumb line.	
IX. General alignment	Note the body as a whole with respect to alignment with the plumb line in place at center of instep.	
X. Bilateral symmetry	Observe from rear the symmetry of head, shoulders, spine, hips, leg; note amount of unevenness if present; turn subject so that plumb line falls down center of back; measure lateral deviation of spine (scoliosis), and observe with subject bent over.	
		Total score:

The chart is scored according to the following key:

Very marked defect needing immediate referral	0–2 points
Marked defect sufficient to be assigned to corrective class	3–4 points
Mild defect sufficient to be assigned to corrective class	5–6 points
Slight defect sufficient for advisement but not corrective class	7–8 points
No defect, perfect posture	9–10 points

Standardized Tests for Evaluating Posture

In using objective instruments to appraise posture the emphasis should be on three things: first, to determine an accurate picture of posture status; second, to relate these findings to health status in general; and third, to use the test results to encourage greater postural consciousness.

PHOTOGRAPHY

Pictures crystallize ideas and form a basis for understanding postural defects, on the part of both the instructor and the pupil. "Before" and "after" posture photographs are helpful. When Fishman started taking pictures with his Polaroid Land Camera, his junior high school pupils came to life and a highly stimulating program of posture appraisal and remedial

Fig. 9-2. The Grid Photograph. (Courtesy Harry R. Fishman, Philadelphia School Department.)

instruction was the result.[29] Pupils were happy to be photographed at ease and to be able to view themselves a minute after the shutter clicked. This is without doubt the psychological moment for optimum health teaching and illustrates once again the close tie between evaluation and instruction.

A small and simple photograph can be made by following a standardized procedure in almost any school. It can be easily inspected and studied through projection by the opaque projector. Markings may be made on it to highlight certain characteristics. Moreover, it is not uncommon to find school systems that place a grid of horizontal and vertical lines behind the subject.[30] Years ago, when film and camera equipment was expensive, the silhouettograph and shadowgraph were employed to give an outline of the torso and highlight certain body landmarks. Today, however, the regularly posed picture, properly related to a grid background and taken with the ordinary still camera and flash equipment, is far superior to the old method, and pupil interest is greater.

Motion pictures are particularly useful in estimating the mechanics of movement in general. Furthermore, the pupil sees himself and helps initiate a program of improvement.

SILHOUETTES

Norman Fradd, at Harvard University, first devised the *silhouettograph* for showing anteroposterior curves. This is a 5- by 7-inch photograph, taken directly on inexpensive bromide paper, indicating the position of the scapulas, abdomen, and spinal curves in profile. When the picture is developed, a white silhouette of the subject appears against a black background. If the backdrop is ruled in 2-inch squares before the picture is taken a white-on-black gridwork will be included in the silhouettograph. By placing dots on the finished picture, lines can be drawn to establish body angles.

This technique is of limited accuracy. Even carefully posed subjects cast a silhouette that is not adequate. Cureton and others have shown that this is due in part to photographic errors of camera adjustment, improper lighting, lack of clear-cut pictures, and misrepresentation of the spinal curves.[31] The silhouettograph, nevertheless has had wide use, both before and after Cureton's research.

Brownell set up a scale for using the silhouette.[32] He had a large group

[29] Harry Ronald Fishman, "One-minute Posture Picture," *Journal of Health, Physical Education and Recreation*, 29:72–74, January, 1958.

[30] Prof. Frank Kleeberger was a pioneer in this technique at the University of Southern California many years before it was used elsewhere.

[31] Thomas K. Cureton, J. S. Wickers, and H. P. Elder, "Reliability and Objectivity of Springfield Postural Measurements," *The Research Quarterly* (Supplement), 6:81–90, May, 1935.

[32] Clifford Lee Brownell, *A Scale for Measuring the Anterio-posterior Posture of Ninth Grade Boys*, Teachers College Contribution to Education no. 325, Bureau of Publications, Teachers College, Columbia University, New York, 1928.

of judges view a random sample of 100 silhouettes and arrange them in order of merit. This was reduced to 13 silhouettes, arranged in rank order and presented in scale scores on a chart. In using the chart the physical education teacher compares the silhouette of the pupil with the chart. This is done by starting from the bottom of the scale and working up; then beginning at the top end of the scale and working down.

Crook, working with preschool children, followed a procedure quite similar to that developed by Brownell; here again posture is rated by thirteen silhouettes.[33] Christenson attempted to improve on Brownell's work by projecting the silhouette of a pupil over the Brownell standard silhouette.[34] This was accomplished by a projection machine. It was demonstrated that judges could agree very closely on how well the silhouette picture of a subject compared with the standard on which it was superimposed.

At about the same time Hubbard devised a photographic procedure employing more light on the subject, which brought out a clearer silhouette picture.[35] Many features of body form and more detail in the area of the lower leg and feet were obtained. This type of picture, although not as clear as a photograph, was cheaper and easy to finish. As with the older silhouettograph, no film or plate is used. Advantages include the following: there is a clear indication of the shoulder blades and direction of the spine; there is improved observation of the muscular development and position of the feet; markings for alignment purposes can be placed on the body before photographing; and the negatives respond very well to enlargement and photostatic reproduction.

Klein examined over 2,200 subjects in order to build a composite silhouette outline norm[36] Korb used this norm as part of a comparograph built to improve the validity of judging silhouettes.[37] By placing an outline of the posture norm on the curtain before which the subject stands, the resulting silhouette of the subject will have an outline of proper posture surrounding it. Employing the opinions of a battery of judges, Korb determined that this technique of analyzing silhouettes increased the validity of this postural measurement by 47.5 per cent.

BUHL-MORRILL POSTUREMETER

An instrument for appraising cervicothoracic posture was developed to make posture measurement more easily understood.[38] Buhl and Morrill tested

[33] Billie Louise Crook, "A Scale for Measuring the Anterio-posterior Posture of Pre-school Child," *The Research Quarterly,* 7:96–101, December, 1936.

[34] Cornell Hjalmar Christenson, "An Improvement in Technique for Measuring Anterio-posterior Posture," *The Research Quarterly,* 4:206–211, December, 1933.

[35] C. H. Hubbard, "Advantages of the New Shadow-silhouettograph over the Original," *The Research Quarterly* (Supplement), 6:50–53, March, 1935.

[36] Armin Klein and Leah C. Thomas, *op. cit.,* p. 55.

[37] Edwin M. Korb, "A Method to Increase the Validity of Measuring Posture," *The Research Quarterly,* 10:142–149, March, 1939.

[38] Thomas K. Cureton, J. S. Wickens, and H. P. Elder, *op. cit.,* p. 81.

1,110 girls between the ages of thirteen and nineteen with a posturemeter, a device consisting chiefly of a protractor, a level to assure that the 90-degree graduation is perpendicular, and a movable finger to correct the anterior point of the horizontal anteroposterior line through the body with a fixed point on the head (the tip of the nose). The basic principle behind the posturemeter is that, as the posture of the thoracic region improves, the chest moves forward and upward and the head moves backward. The test is limited to only one section of the body, and the instrument is difficult to obtain. The principle of body-angle measurement, however, has value if the procedure is not too complicated for school use.

CURETON-GUNBY CONFORMATEUR

Because of the difficulty of scaling photographs and silhouettographs, Cureton set up a device called the conformateur. The anteroposterior spinal curves are measured directly in full life size on conformateur rods, which project horizontally from a vertical upright. The metal rods or spindles slide horizontally through holes bored into the vertical upright. Clamps hold each rod in place. The person to be measured stands with the back to the rods so when the rods are tapped lightly they will make contact with the spinous process of the spine, from sacrum to the top of the head. The conformateur can also be extended down the body to the heels. By locking the rods in place a highly accurate measurement of spinal deviations is made. Cureton finds that the objectivity is about two to four times as good as subjective judgments repeated a week apart.

Norms are available on photostatic sheets.[39] They are for young men only. About an hour is needed to measure and scale each subject. The value of the conformateur is limited to good equipment and careful training of the anthropometers. Cureton has not related the test results to aspects of fitness, but as a measurement of the anteroposterior spinal curves the device has value.

WELLESLEY POSTURE TEST

The staff at Wellesley College worked with MacEwan and Howe to construct a posture test for women that would measure the degree of curvature in the dorsal and lumbar spine, body tilt and segmental angulation, and the position of the head and neck.[40] This was done from photographs of the standing subject. To make appraisal more objective eleven aluminum pointers, 9 centimeters long, 4 millimeters wide, and 0.25 gram in weight, were attached to the subject's skin by narrow strips of adhesive tape. They

[39] See *The Research Quarterly* (Supplement), 12:356, May, 1941.

[40] C. C. MacEwan and E. C. Howe, "An Objective Method of Grading Posture," *The Research Quarterly*, 3:144–147, October, 1932.

were attached at the end of the sternum, on the prominence of the first piece of the sacrum, and on the spinous processes of every other vertebra starting with the seventh cervical. After the subject has been photographed, the true spine and chest position are easily observed. MacEwan and Howe set up a numerical scale for grading posture 1 to 25; this could be changed to a letter grade of A+ to E—.

The test was validated by eight judges experienced in the area of female body mechanics. Its use, in the present form, however, is questionable. Much time is used in preparing the subject for photographing and in scaling and rating the picture later. The idea of pointers to highlight body landmarks is a good one.

WICKENS-KIPHUTH POSTURE TEST

Since the beginning of the corrective exercise programs at Yale University, in 1919, a photograph has been taken of the side view of students. This becomes a part of the permanent personnel record of the student and over the years has proved valuable in the study of body mechanics.

Using aluminum pointers similar to those used at Wellesley College, Wickens and Kiphuth, in 1937, attempted to make more objective appraisals of the standing posture of freshman students.[41] The pointers were placed on the spinous process of the seventh cervical vertebra, at the point of the greatest convexity in the upper back, at the point of inflection between the dorsal and lumbar curves, at the point of greatest concavity in the lower back, and on the prominence of the sacrum. In addition, flesh pencil marks were placed on the left side of the subject in order to determine the antero-posterior alignment. These marks were placed at the tragus of the ear, the tip of the acromion, the greater trochanter, the head of the fibula, and the cuboid bone of the foot. A photograph was then taken from the left side of the subject, and from this photograph definite angles in the upper and lower back could be noted, as well as the position of the head and neck in relation to the body as a whole, and the over-all tilt of the body in the anteroposterior position.

In 1952, Photo Metric equipment was installed at Yale University to provide four images of the subject in one exposure (front, rear, side, and overhead views). It is possible, from detailed engineering studies, to obtain accurate measurements of any part of the body.[42] A slide, made of each exposure, is projected half life size on the screen. Various desired measurements, too small to be taken by photographic prints, are taken with reasonable accuracy. When the slide is projected, marks are placed on the screen

[41] Stewart Wickens and Oscar W. Kiphuth, "Body Mechanics Analysis of Yale University Freshmen," *The Research Quarterly,* 8:38–44, December, 1937.

[42] T. Erwin Blesh, Carlton R. Meyers, and Oscar W. Kiphuth, *Photometric Photography in Posture Evaluation of Yale University Freshmen,* Yale University Clerical Bureau, New Haven, Conn., 1954.

itself to indicate where the ends of the pointers actually fall on the body. When all points are located on the screen, measurement is begun on the side view with a specially designed protractor-ruler to determine the angles that have been used in the Yale program in the past.

The fundamental measurements recorded are as follows:[43]

1. Head and neck position (see Figure 9-3, Angle A): The position of the head and neck in relation to the trunk is determined by scaling the angle formed

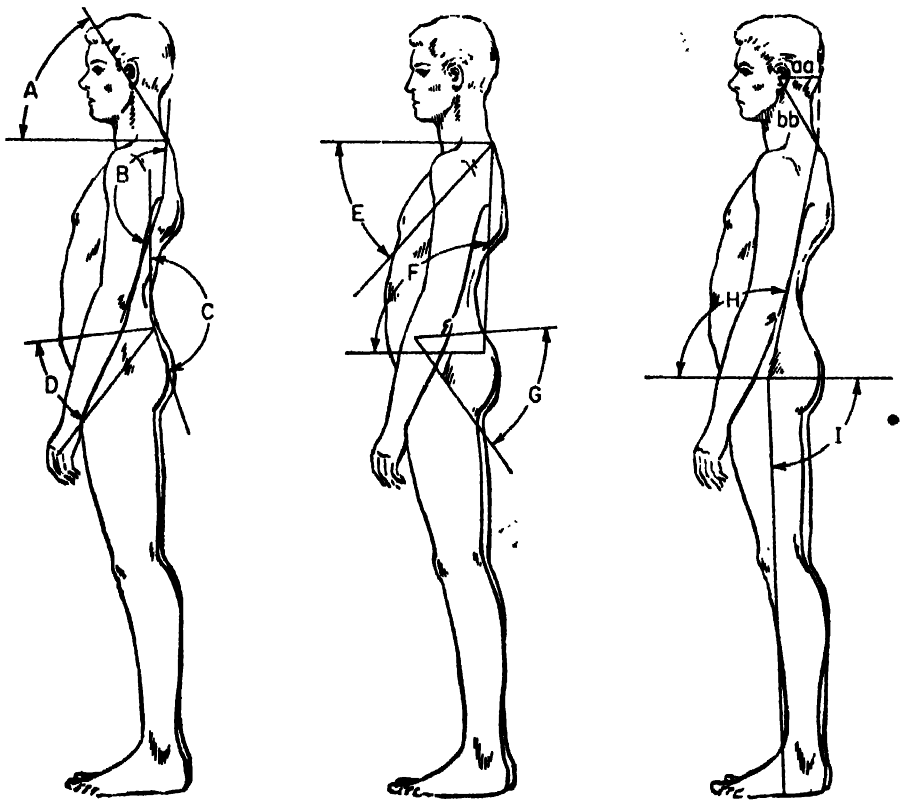


Fig. 9-3. Fig. 9-4. Fig. 9-5

Figs. 9-3, 9-4, 9-5. The Yale University Wickers-Kiphuth Posture Test. (Courtesy Yale University.)

by a horizontal line running through the seventh cervical vertebra and a line drawn from that vertebra through the tragus of the ear. The size of this angle will vary as the head is thrust forward and backward, and the more the head is thrust forward the smaller the angle will be.

2. Kyphosis (see Figure 9-3, Angle B): The amount of curvature of the upper back is found by scaling the angle which is formed by a line drawn through the point of greatest convexity and the seventh cervical vertebra and a line from the

[43] *Ibid.*, pp. 9–12.

point of greatest convexity through the point of inflection between the dorsal and lumber curves. The greater the roundness of the upper back, the smaller the angle will be.

3. Lordosis (see Figure 9-3, Angle C): The curve in the lower back is determined by scaling the angle formed by a line through the greatest concavity and the point of inflection and a line between the point of greatest concavity and the prominence of the sacrum. This angle of concavity will become smaller as the hollowness becomes greater.

4. Overcarriage (see Figure 9-4, Angle F): The backward tilt of the trunk in relation to the hips is ascertained by measuring the angle formed by a horizontal line through the prominence of the sacrum and a line from that point through the seventh cervical vertebra. As the trunk begins to tilt back over the hips this angle will approach, or surpass, a right angle.

5. Chest (see Figure 9-4, Angle E): The position in which the chest is carried is disclosed by scaling the angle formed by a horizontal line through the seventh cervical vertebra and a line drawn from the end of the sternum to the seventh cervical vertebra. The higher the chest is carried, the smaller the angle will be.

6. Hip thrust (see Figure 9-5, Angle I): The position of the hips in relation to the feet is determined by measuring the angle formed by a line drawn from the cuboid bone through the greater trochanter and a horizontal line through the greater trochanter. As the hips are thrust forward, the angle becomes smaller.

7. Linear measurements (see Figure 9-5, Angles aa and bb): The half life-size image makes it practical to measure certain linear distances, (*a*) the horizontal displacement of the tragus of the ear from the seventh cervical vertebra (angle aa) and (*b*) the linear distance from the tragus of the ear to the seventh cervical vertebra (angle bb).

8. Pelvic tilt (see Figure 9-4, Angle G, and Figure 9-3, Angle D): Two angles are measured on the anteroposterior view to determine the more significant indicant of pelvic tilt. The reference points on the left side for both angles are the anterior superior spine of the ilium, posterior superior spine of the ilium, and the greater trochanter of the femur. The anterior superior spine of the ilium serves as the vertex for the first angle (Angle G), while the posterior superior spine is the vertex for the second angle (Angle D).

Other measurements can be made of the knees and shoulders. Shoulder displacement is especially easy to record accurately with the overhead picture. An overcarriage angle (Angle H, Figure 9-5) is also made by a horizontal line through the greater trochanter and a line from the latter to the seventh cervical vertebra. Trunk and leg relationships can be compared this way.

By and large, the Photo Metric technique is a superior and highly practical method of measuring postural angles. It provides sound data for a remedial physical education program. Moreover, when the special equipment is not available, a photograph can be taken following the same procedures. With the opaque projector the photograph can be projected for measurements, and silhouettes can also be used.

WOODRUFF BODY ALIGNMENT POSTURE TEST

In an effort to get away from the need for posture photographs, Woodruff, at the University of Oregon, built a posture frame.[44] The frame stands on a base just over 6 feet high. It is 9 inches wide and has nine strings ¾ inch apart running from top to bottom in the frame. For use the frame is placed approximately 6 feet from a wall on a base line running from the wall at a right angle. Another line is drawn on the wall perpendicular to the floor so it intersects the floor line. The subject to be appraised stands on a spot between the wall and the alignment frame. The examiner stands on a spot 10 feet in front of the frame, looks through the frame, and lines the center string of the frame with the wall line. The plumb-line measurement can be noted in this manner. Deviations in posture are scored by noting the deviation, either forward or backward, for the number of strings at each body part; ankle, knee, hip, shoulder, and hand. The perfect score is 25; points are deducted for each string deviation. The range of scores is from 16 to 25, with a mean of about 20.[45]

HOWLAND ALIGNOMETER

A technique for teaching body alignment in standing was developed by Howland; it consists of the perpendicular alignment of two anatomical landmarks on the body trunk: the center of the sternum and the superior border of the symphysis pubis.[46] It was measured against the traditional method of aligning five anatomical landmarks (lobe of ear, acromion process, greater trochanter, knee, and ankle joints) as the criterion for body alignment in standing. It was further validated by the use of radiographs, photographs, and a measuring instrument, the Alignometer, which was developed by Howland.

The Alignometer is simple in construction and not difficult to make. It consists of an easy arrangement of two sliding calibrated pointers affixed to a perpendicular steel rod. The rod is supported on a wooden base (see Figure 9-6). The varying heights of subjects tested can be allowed for by sliding the arms that support the pointers up and down. By the use of a vertical calibrated rod connecting pointers A and C (Figure 9-6), it is possible to determine the exact center of the sternum, that is, the mid-point between the top of the sternum (pointer A) and the bottom of the sternum (pointer C). When this is determined, the set screw on the arm (pointer B) is tightened. Likewise the lower pointer (D) is brought out just far enough to make contact with the upper border of the symphysis pubis. If the subject

[44] Janet Woodruff, School of Health and Physical Education, University of Oregon, Eugene, Ore.

[45] H. Harrison Clarke, *The Application of Measurement to Health, Physical Education, and Recreation,* Englewood Cliffs, N.J., Prentice-Hall, Inc., 1959, p. 157.

[46] Ivaclare Sprow Howland, *op. cit.*

is properly aligned on a vertical plane the two pointers (B and D) will be vertically in line; that is, each pointer will be the same distance from the vertical rod.

In Howland's research a rectangular parallelogram between the sterno-pubic alignment and the reference line (plumb line) of the body occurred in each test. Thus the upper and lower portions of the body were aligned when there was a favorable reading on the Alignometer. Objectivity is approximately .90 and reliability .92. There is little reason why this instrument should not have widespread use with children and adults of all sizes. Especially noteworthy is its value in measuring the extremes of pelvic tilt and in teaching subjects the proper position to maintain for improvement. Howland's greatest success has come from using the instrument as a teaching device in body mechanics classes and by teachers of speech, voice, and dramatics interested in poise and body control.

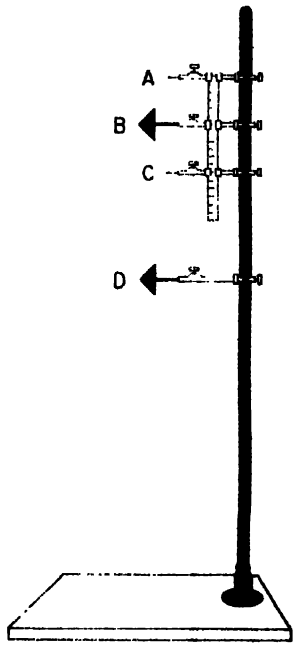


Fig. 9-6. The Howland Alignometer. (A: Pointer to locate top of sternum. B: Pointer to locate center of sternum. C: Pointer for bottom of sternum. D: Pointer to locate symphysis pubis.) (Courtesy I. S. Howland.)

NEW YORK STATE POSTURE RATING TEST

A major part of the New York State Physical Fitness Test is devoted to the appraisal of posture. A series of thirteen posture profiles appears on the score sheet. Each profile is shown in three degrees of variation (Figure 9-7): the proper position (5 points); the slightly improper position (3) points); the markedly improper position (1 point). The profiles cover the head, shoulder position, shoulder level, spine, hip level, feet, arches, neck, chest, upper back, trunk, abdomen, and lower back. The points for each item are entered under the proper school grade (four to twelve) in the scoring column. All thirteen scores are totaled for the subject's grade level and put in the appropriate place for evaluation of total fitness.

Although appraising posture by this kind of observation is relatively subjective, there is nevertheless some value in it. In New York State, where the use of the test is widespread, the skill of the physical education teacher appears to improve considerably as more and more pupils are appraised. Moreover, those pupils with a low posture score are frequently referred to the school health service for a special physical examination. As with all physical fitness appraisal measures, the cooperation of the school physician and his staff is most desirable.

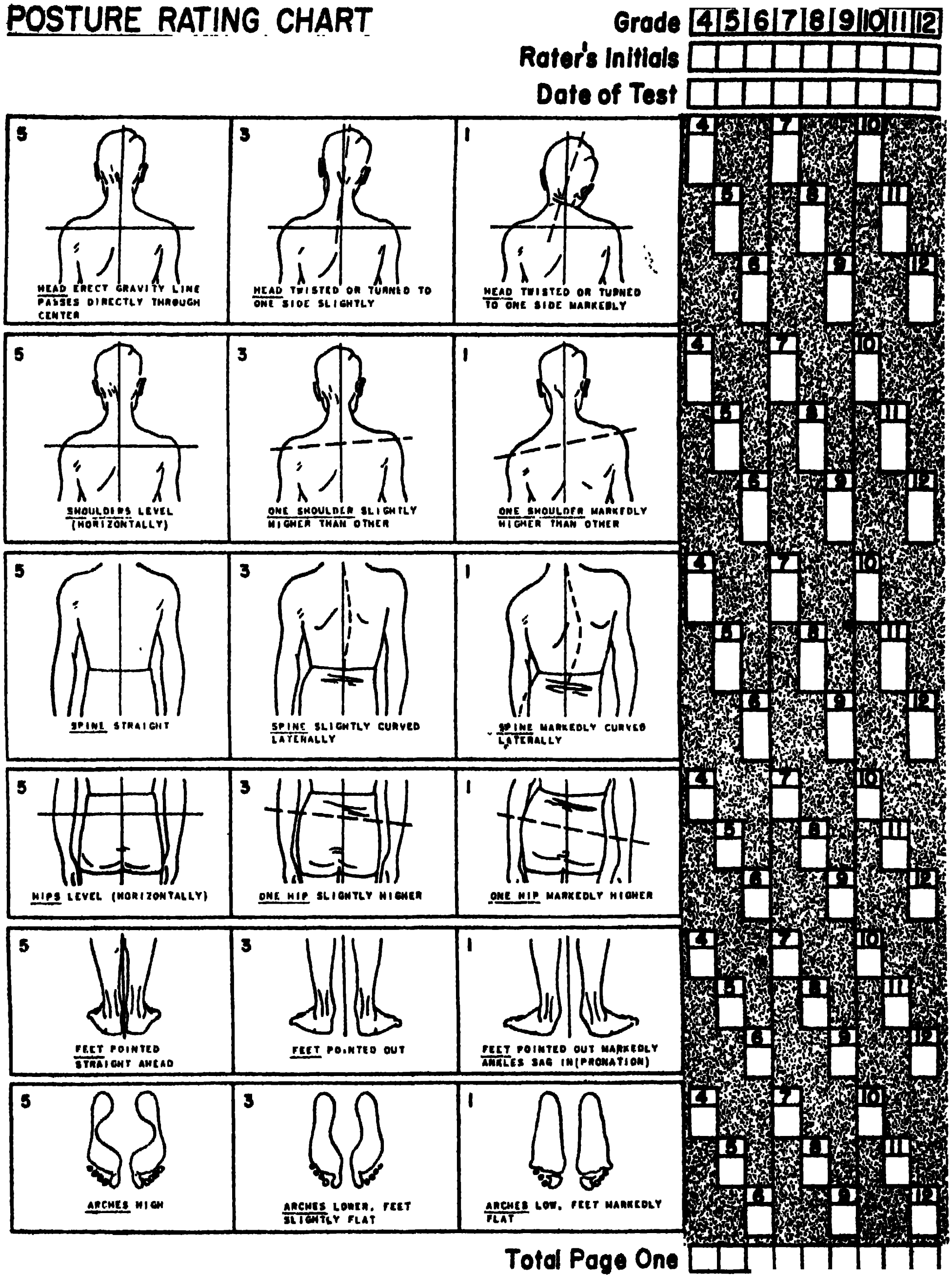
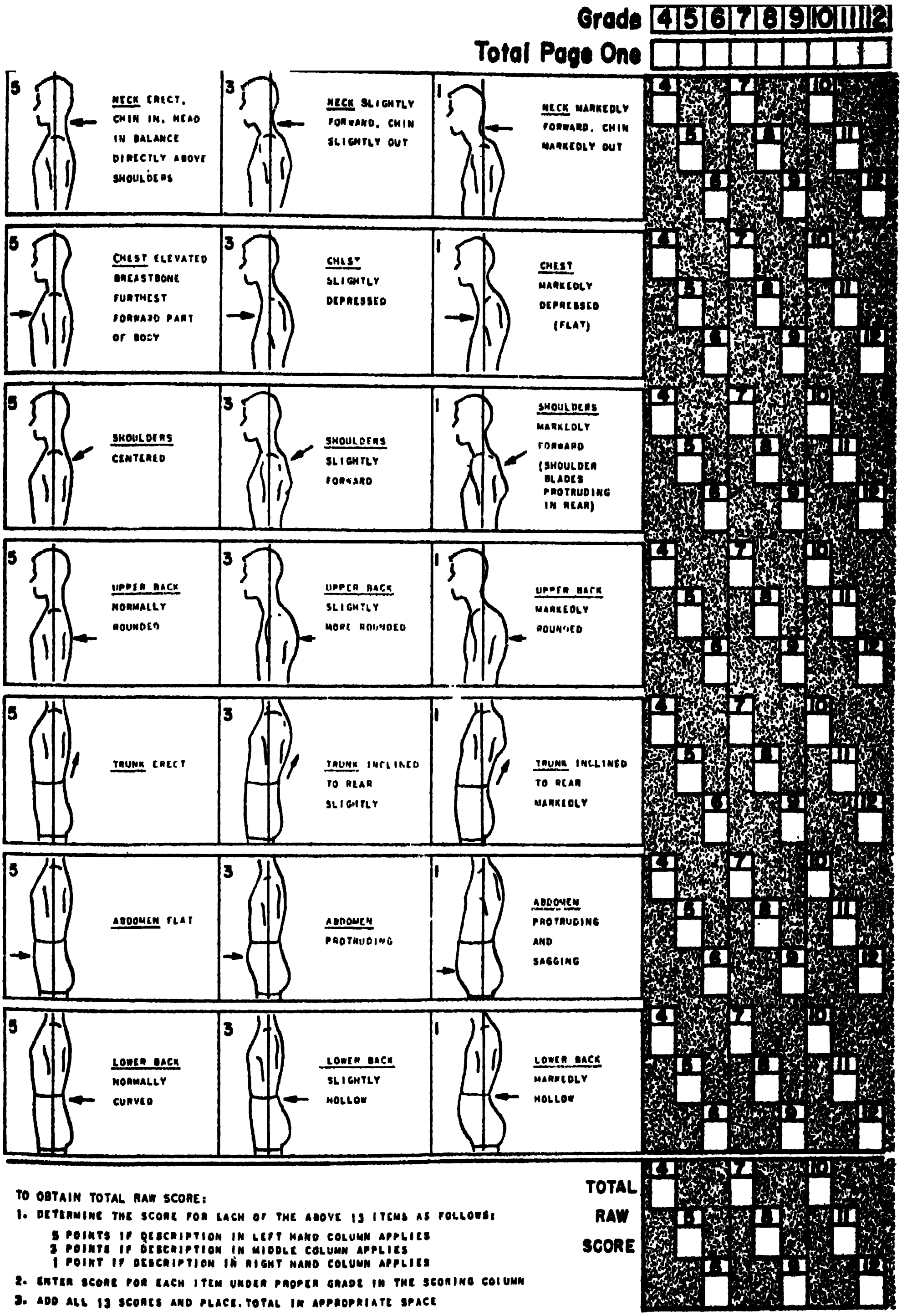


Fig. 9-7. A Section of Posture Profiles from the New York State Phy-



sical Fitness Test. (Courtesy New York State Education Department.)

MASSEY POSTURE TEST

Massey selected three qualified judges to set up a criterion of good anteroposterior posture.[47] He then carefully silhouetteographed 200 male students and proceeded to develop a series of angles and indices that could be measured from the silhouette. The angles selected for posture measurement correlated .985 with the established criterion. The angles were as follows: Angle I, head-neck-trunk alignment; Angle II, trunk-hip alignment; Angle III, hip-thigh-knee alignment; and Angle IV, thigh-leg-ankle alignment. The test correlates quite well with the old Kellogg battery (pelvic obliquity, chest ratio, and head angle) that was in vogue about 1900.

CENTER OF GRAVITY TEST

Much has been said about the vertical or gravity line as it relates to anteroposterior posture. Cureton and Wickens examined a number of subjects to see how they stood with reference to the line of gravity.[48] This finding was then related to such items as general physical fitness, athletic ability, and posture. In order to do this it was necessary to set up carefully a balance board on two scales of similar calibration. A scale was placed at each end of the balance board, and the subject was asked to stand exactly in the middle of the board so that the internal malleoli could be lined up with a vertical pin located exactly in the center of the board. Toledo scales were used. Each scale was read and the reading recorded. Half the weight of the balance board is subtracted from the reading of each scale. By comparing the scale scores it is possible to show how far the line of gravity has fallen in front or behind the internal malleoli. Norm tables have been prepared so a percentile score can be calculated.

This work is not difficult to repeat, and although it seems a little crude, it is nevertheless quite accurate in noting the forward lean (ahead of the gravitational line) so often found in athletes and persons of mesomorphic body build. In this respect Cureton and Wickens discovered a considerable degree of body lean and flat upper backs in persons of high physical fitness. Athletic ability, as demonstrated by the Sargent Vertical Jump, correlated .49 with the center of gravity test and a more forward center of gravity. The athlete appears to be ready to go and is not in a position to be caught "flat-footed." The test also correlated .50 with the Rogers Strength Index and as high as .75 with the Physical Fitness Index. In kyphosis, when the gravity

[47] Wayne W. Massey, "A Critical Study of Objective Methods for Measuring Anterior Posterior Posture with a Simplified Technique," *The Research Quarterly,* 14:3–10, March, 1943.

[48] Thomas K. Cureton, and J. Stuart Wickens, "The Center of Gravity Test of the Human Body in the Anterio-posterior Plane and Its Relation to Posture, Physical Fitness and Athletic Ability," *The Research Quarterly* (Supplement), 6:93–100, May, 1935.

line shifts to compensate for the cervical curve, the correlation with the test was —.36.

CLARKE-SHAY SCOLIOMETER

Few tests have been standardized for noting the degree of lateral curvature of the spine (scoliosis). The curve can frequently be observed, and even photographed, but the degree of severity or improvement that follows remedial work is often difficult to appraise accurately.

Such devices as the Taut String Test have been employed to note the lateral curvature. In this case a string is held taut between the spinous processes of the seventh cervical and fifth lumbar vertebrae. By observing the amount of lateral deviation of the vertebrae between the points where the string is held it is possible to estimate the severity of the defect. As might be expected, it is hard to hold the string tight, and there is no way to measure objectively the degree of deviation from the vertical line.

In an effort to remedy the situation partially, Clarke and Shay devised a simple scoliometer made of transparent celluloid.[49] The celluloid, long enough to cover the spinal area and 16 centimeters wide, is ruled with longitudinal gradations 4 centimeters apart. The subject to be tested is marked with a skin pencil on the tops of all spinous processes of the vertebrae, plus the spines and lower angles of the scapulas. After the marking the celluloid scoliometer is placed over the spine so any lateral curve will show through and the degree of deviation is read on the longitudinal scale. Difficulties involving the twisting of the celluloid and the fact that a subject does not always assume quite the same position have tended to reduce the value of the scoliometer.

KRAUS-WEBER REFINED POSTURE TEST

Over many years Kraus and his associates worked with children and adults in need of posture improvement. Numerous experimental measurements were taken on the body, and a number of basic rules for measuring posture were developed.[50] A series of structural measurements using body landmarks was set up. It included chest expansion, scapula-spine distances, level of the scapulas, level of the anteroposterior spine of the ilium, length of the legs, angle of pelvic tilt, dorsal kyphosis, and lumbar lordosis. Also there were a number of functional measurements to test the elasticity and strength of muscle groups. More specifically tests were established for testing the elasticity of the pectoral muscles, hamstring tendons, and erector spinae muscle plus hamstring tendons.

[49] H. Harrison Clarke and Clayton T. Shay, "Measurement of Lateral Spinal Deviations," *Black and Gold of Phi Epsilon Kappa,* 17:38–40, March, 1940.
[50] Hans Kraus and S. Weber, *op. cit.*

Evaluating the Feet

The structural and functional status of the foot should be included in any study of posture. Feet examined and well-cared for during the formative stages of development will, in most cases, function properly during middle age, when persons become heavier and less active. Poor feet at the base of the human structure tend to disrupt body alignment and create problems in the upper extremities. Appraising the efficiency of the feet of school children has value, providing the process is not too time-consuming and complicated.

From the standpoint of routine class observation the feet can be checked by the teacher in several ways. The occurrence of one or more of the following items is cause for a more careful check of foot health by the school physician or specialist:

Irregular gait in walking.
Arches of the feet are flat.
Feet toe out too far in standing and walking.
Achilles tendon above heel bone is bowed inward when standing.
Ankles turn in.
Heels of shoes wear down on inside borders.
Pain in the bones of the foot or ankle.
Pain in the lower legs.

Needless to say, some pupils who appear to have a foot difficulty may have a perfectly normal foot. The slightly bowed ankle does not bother some persons, and the congenital flat foot is frequently confused with the fallen longitudinal arch. Because of this it is good practice to single out suspected cases of foot trouble and investigate further the functional efficiency of the foot. This can be accomplished in school by questioning the pupil, by seeking pain in the foot, and by careful observation of his performance in walking, running, and jumping.

KELLY FOOT PAIN TEST

Kelly developed a test noting foot pain in children with symptoms typical of functional foot strain.[51] In this test pain on the sole of the foot due to applied pressure differentiates normal from painful feet as well as pronated from painful feet. Kelly's test uses a bathroom scale on top of which rests a small platform with a ½-inch diameter padded nailhead. By pressing each of three pressure points of the foot on the nailhead the subject can note the presence of pain. If no pain occurs up to 12 pounds pressure on the scale the foot is not considered weak. The pressure points are (1) under the junction of the first and second metatarsals with the first cuneiform bone, (2) under the insertion of the tibialis anterior muscle, and (3) under the

[51] Ellen D. Kelly, "A Comparative Study of Structure and Function of Normal, Pronated, and Painful Feet among Children," *The Research Quarterly,* 18:291–295, December, 1947.

posterior insertion of the plantercalcaneo-navicular ligament into the calcaneous bone.

THE FOOTPRINT ANGLE

Schwartz originated the footprint angle in the U.S. Public Health Service, in 1928.[52] He demonstrated that the angle of the arch increases with the height of the arch. Both Clarke and Cureton used the original footprint angle and obtained objectivity and reliability coefficients of .95 to .97. Cureton proved the validity of the footprint angle as a measure of the height of the arch by using sand impressions.[53] He did not find a significant correlation between the high arch and functional performance. Later, however, Cureton obtained a correlation of .613 between height achieved in the vertical jump and four variables: (1) internal malleoli height, (2) toe flexion strength, (3) ankle extension strength, and (4) center of gravity.[54] This is not a high correlation, but it is of some value when one considers that the vertical jump correlates fairly well with such activities as the running high and broad jump.

For many years a machine called the pedograph was used to record the footprint and angle on a standardized score sheet. A permanent record was available, and improvements could be noted.

Clarke improved the objectivity of the original footprint angle test.[55] The test is as follows:

a. Place a piece of heavy brown paper on the floor beside a damp towel.

b. Have the subject step on the towel with the right foot and then step on the brown paper.

c. Before the footprint dries on the paper draw lines A–B and C–D as indicated in Figure 9-8.

1. Draw a line to represent the medial border of the foot between the points of the imprint at the base of the first metatarsal bone (base of the big toe) and the calcaneous or heel bone.

2. Locate a point where this line first touches the imprint on the inner side of the big toe.

3. Then hold a ruler on this point and swing it down from the toes until it just touches the edge of the print on the inside of the arch and draw a line from the point across the print. No brown paper should show between this line and the print.

[52] Louis Schwartz, R. H. Britten, and L. R. Thompson, *Studies on Physical Development and Posture,* U.S. Public Health Service Bulletin, no. 179, 1928, p. 23.

[53] Thomas K. Cureton, "The Validity of Footprints as a Measure of Vertical Height of the Arch and Functional Efficiency of the Foot," *The Research Quarterly* (Supplement), 6:70–80, May, 1935.

[54] Thomas K. Cureton, "Fitness of the Feet and Legs," *The Research Quarterly* (Supplement), 12:368–380, May, 1941.

[55] H. Harrison Clarke, "An Objective Method of Measuring the Height of the Longitudinal Arch of the Foot," *The Research Quarterly,* 4:99–103, October, 1933; Permission obtained to reproduce the measuring instruction accompanying Figure 9-8.

d. Measure the footprint angle at the junction of the two lines with a protractor and record.

A satisfactory angle is about 42 degrees. Feet with angles of less than 30 degrees should probably be considered for further examination and correction. Unless the arch is a normally flat one, it will usually respond to special foot exercises; footprint angle measurements taken at a later date will show an improvement.

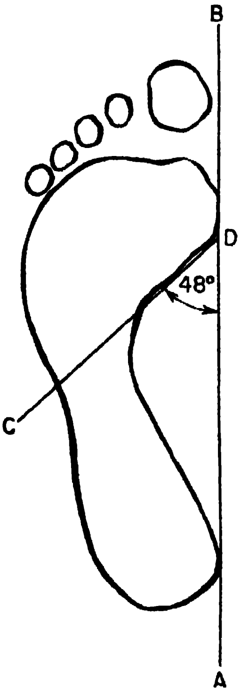


Fig. 9-8. Clarke Footprint Angle.

PEDORULE

This instrument is a rectangle of heavy plate glass, 7 inches wide and 9 inches high, with the surface parallel-lined in tenths of inches.[56] The pedorule was devised to measure the amount of deflection of the tendon of Achilles from the perpendicular. In the normal foot the tendon of Achilles makes a straight line, but in a weak foot the lower end tends to be deflected outward. Before measuring a person, two ink dots must be placed on the tendon: the mid-point as high on the leg toward the calf as possible and the mid-point of the back of the heel. When the subject is measured, the center line (a heavier line) should then bisect the two ink marks on a normal foot.

The most satisfactory method of using this instrument is to place the center line of the pedorule directly behind the center of the tendon at the point where it is bowed inward farthest. The distance that the tendon deviates from the perpendicular is easily found by counting the number of lines from the point of maximum bowing to the internal malleolus. This figure is subtracted from the distance from the center of the bowed-in tendon to the external malleolus. The perfect foot, of course, would be zero. Test reliability is .94.

Danford feels that the pedorule is superior to the footprint angle as a test for foot efficiency. The low arch may not bother a person, but the bowed tendon frequently causes painful feet.[57] Moreover, Danford demonstrated that even with a normal arch, around 42 degrees, it is possible to have a bowed tendon and painfully inefficient feet. His findings substantiate those of Rogers that the correlation of agreement between foot measurements indicate only .28 between left and right feet. Therefore, both feet should be appraised.

[56] Harold R. Danford, "A Comparative Study of Three Methods of Measuring Flat and Weak Feet," *The Research Quarterly* (Supplement), 6:43–49, March, 1935.
[57] *Ibid.*, p. 48.

Posture Evaluation in Remedial Physical Education

Generally speaking, the school or college with a remedial program of physical education is doing the most for pupils in lowered states of fitness. It is in these programs, often carried out on a cooperative basis with the school physician and nurse, that posture appraisal and correlation usually takes place. Because so many variables (nutritional status, mental and emotional health, disease, etc.) influence a case of poor posture, it cannot be emphasized too strongly that test results need careful interpretation. This is all the more true because a number of persons with relatively poor posture appear to function adequately in their day-to-day activities. By and large, however, poor posture signals that all is not perfect: physically, socially, or esthetically.

There is evidence that across the nation boys and girls are being screened for poor posture; sometimes this is a part of a yearly physical examination, and at other times it is part of general physical education. In some communities the posture contest is employed regularly to focus attention on physical development. In Alhambra, California, for example, all students are screened for posture twice a year. The pupils are scored, and class winners compete in a posture contest. From the semifinals ten contestants are chosen and awarded a printed certificate in body mechanics By the time of the final contest there is great community interest. Community physicians, therapists, and educators from local universities assist in the final judging. Contestants are scored on sitting, walking, certain exercises, posture screening, plumb-line test, foot and leg alignment, general body structure, and the opinion of roaming judges who evaluate posture as students pass between stations. The total effort does much to promote posture in the minds of the public.

In an extensive survey of physical education for children of elementary school age, the U.S. Office of Education found that 24 per cent of the school systems reporting gave information on the methods they use to evaluate children's posture.[58] The most common methods employed were check lists or tests at regular intervals, classroom teacher's subjective judgments, periodic health examination, Bancroft Posture Test at varying intervals, and posture examinations given by school nurse. Several other methods were also mentioned in individual reports. The report of one particular school system is most impressive, for it makes posture appraisal one of the important tasks a school can perform for its pupils. It is full of meaning and is worth repeating here[59]:

A notice is sent home that posture pictures will be shown during the physical education period. Parents are urged to come to see the pictures and to learn about

[58] *Physical Education in Urban Elementary Schools*, U.S. Department of Health, Education and Welfare, Office of Education, Bulletin no. 15, 1959.

[59] *Ibid.*, p. 44.

the posture of their children. The examiner talks to each group, pointing out things that can be corrected and recommending exercises. While this is being done, the other examiner has taken the equipment and two clerks to another school. Last year over 10,000 children were examined. Parents turn out in large numbers for the picture and for the conferences.

SELECTED REFERENCES

Broer, Marion R.: *Efficiency of Human Movement,* Philadelphia, W. B. Saunders Company, 1960, chaps. 1 and 2.

Byrd, Oliver E.: "School Health," *Childhood Education,* 35:394–397, May, 1959.

Fox, Margaret G.: "The Relationship of Abdominal Strength to Selected Posture Faults," *The Research Quarterly,* 22:141–144, May, 1951.

Gurewitch, A. D., and M. O'Neil: "Flexibility of Healthy Children," *Archives of Physical Therapy,* 25:216, April, 1944.

Kelly, Ellen D.: "A Comparative Study of Structure and Function of Normal Pronated, and Painful Feet among Children," *The Research Quarterly,* 18:291–295, December, 1947.

Kraus, Hans, and S. Weber: "Evaluation of Posture Based on Structural and Functional Measurements," *Physiotherapy Review,* 25:267–275, November–December, 1945.

——— and ———: "Quantitative Tabulation of Posture Evaluation: Based on Structural and Functional Measurements," *Physiotherapy Review,* 26:235, September–October, 1946.

Lowman, Charles Leroy, and Carl Haven Young: *Postural Fitness,* Philadelphia, Lea & Febiger, 1960, chap. 6.

McCloy, Charles H.: "X-ray Studies of Innate Differences of Straight and Curved Spines," *The Research Quarterly,* 9:50–57, March, 1938.

Morton, Dudley J., and Dudley D. Fuller: *Body Locomotion and Human Form,* Baltimore, The Williams & Wilkins Company, 1952.

Phelps, W. W., R. J. H. Kiphuth, and C. W. Goff: *The Diagnosis and Treatment of Postural Defects,* 2d ed., Springfield, Ill., Charles C Thomas, Publisher, 1956.

Steindler, Arthur: *Kinesiology of the Human Body under Normal and Pathological Conditions,* Springfield, Ill., Charles C Thomas, Publisher, 1955.

CHAPTER 10

Tests of Skill and Achievement

To be physically educated means to be able to use one's body efficiently. Education primarily through large muscle activity involves a number of refined skills. Usually several are combined to permit the motor coordination necessary for individual and team game participation. In addition to being the backbone of physical education, skills acquired during the early years relate to fitness in the later years. In fact, all too often the degree of recreation engaged in is directly proportional to the level of physical skill one possesses for a given activity. Tests that measure specific skills, therefore, are fundamental to physical education.

Centuries ago, Aristotle rated motor skills with music and art as necessary items in making the man. Skill tests, when used to measure fundamental movements of walking, running, jumping, climbing, dodging, swimming, handling simple tools, and using pieces of equipment, measure motor processes vital to life. Neuromuscular skills not only have a distinct bearing on general physical behavior, but they contribute to the social efficiency of boys and girls. Numerous studies by men like Gesell, Hardy, Cowell, Jones, Rarick, McGraw, Keister, and Hale show that children superior in gross motor skills are better adjusted than are children who exhibit poor ability in these skills.[1]

The place of sports skills in western civilization is ably defended by Cozens and Stumpf in their inspiring treatment of sports on the American scene.[2] It is an admirable treatise, which attests that Americans are not exclusively sports spectators but are first-degree participants when they have the abilities. Moreover knowledge and ability in a sport produces the best kind of spectator or fan. He becomes a bright, sophisticated observer, who expresses his judgment as a result of his understanding and appreciation. Both the participant and the spectator alike seek excellence. Both are interested not alone in scores, reputations, and action but in *quality of performance.*

[1] See especially the work by Harold E. Jones, *Motor Performance and Growth*, Berkeley, Calif., University of California Press, 1949, pp. 160-170.

[2] Frederick W. Cozens and Florence S. Stumpf, *Sports in American Life*, Chicago, University of Chicago Press, 1953.

Use of Skill Tests

The skill test has its greatest value when it is used to improve instruction. It is concerned with the level of performance. This is just as true in the primary grades with a simple skill, such as dodging or hopping, as it is at the secondary level with a complex sports skill.

Proficiency in a skill demands the nerve-muscle functions of agility, speed, poise, rhythm, and precision of movement. These items, when applied to a specific skill, produce the athlete and dancer, both of which demonstrate a remarkable gracefulness in movement. This quality is a source of inspiration for others. Moreover, when a pupil masters one or two basic skills he is often motivated to learn the more complicated actions. It is not uncommon for school departments to make pin and certificate awards for the successful accomplishment of sports skills. This is especially true in swimming.

Through skill testing, school achievement records may be made more meaningful, especially when it is time for grading. Prospective athletes may be discovered in a physical education class. Also, intramural teams can be equalized in an effort to set up a fair situation for competitive play.

The Validity of Skill Tests

Simply stated, a test is valid if it actually measures what it is supposed to measure. For example, a test for soccer ability is valid if the subject who scores well on it also plays well in the game. Frequently, however, a person scores satisfactorily on the elements of a skill or combination of skills only to play poorly. This sometimes happens on the local level with "homemade" tests. Moreover, it is not difficult for such a condition to arise; a pupil may show considerable skill in several of the separate elements of a sport but lack the ability to combine them properly in a competitive position.

In constructing tests of skill one should know the skill activity well. For soccer, one should understand the game, its objectives, fine points of strategy, equipment, aspects of team play, etc. With this as a background, the game is studied, and the pertinent skills and groups of skills are listed. It may be necessary to try out several combinations of soccer skills involving dribbling, passing, heading, trapping, etc., before a test is of value in selecting the better team players. Furthermore, since it is quite difficult to measure accurately soccer playing ability as demonstrated during a game, it may be advisable to use a committee of competent soccer judges to select the best players. In such a case the test criterion of validity is based on the judgment of experts. Standardized tests of skill are frequently made along these lines. Stroup, working with basketball skills, found that over the years judgment ratings have been the predominant means of validating performance tests.[8] Also, he found that his battery of skills tests predicted the outcome of approximately 84 per cent of the games.

[8] Francis Stroup, "Game Results as a Criterion for Validating Basketball Skill Test," *The Research Quarterly*, 26:353–357, October, 1955.

Another factor to consider in constructing skill tests is the number of trials of a skill movement that are permitted. Frequently a person does much better with more than one trial. Broer discovered that reliable scores in a volleyball volleying test and a basketball-handling test were obtained with three trials.[4] Also, she found that four or five trials are necessary for a reliable score on the basketball-shooting and softball-batting tests.

Most useful tests are subject to some limitations. For example, a test may not possess a high enough validity coefficient to be used for predictive purposes, but it may have instructional value for the teacher. The following are examples of measures that have proved useful over the years; note their limitations.[5]

Test	*Limitation*
Bassett-Glassow-Locke Volleyball Skill Test	Limited to college women
Bennet Test of Diving	Limited to beginning classes
Buchanan Speedball Test	High school girls only; some items have low validity
Dyer Backboard Test of Tennis Ability	Advanced players will profit from a more distant restraining line
Dyer, Schurig, and Apgar Basketball Test	Limited to women; time-consuming
Johnson Basketball Test for Men	No norms; much equipment needed
Heath-Rodgers Soccer Test	Data on elementary children only; reliability and validity low
French Badminton Test	Test may be limited to ten trials on each part to save time
Knox Basketball Test	More useful with experienced players; no norms
Rodgers-Heath Softball Test	T scale limited to fifth and sixth grades
Russell-Lange Volleyball Test for Girls	Data on girls only

Specific Measures of Skill and Achievement

Frequently a number of skills are placed on a separate chart or scale for purposes of motivation or to simplify appraisal by the teacher. For example, at each elementary grade level a pupil should be making some progress in physical skills. An achievement chart hanging on the schoolroom wall might show the degree of advancement in skills such as speed in running, push-ups, rope climbing, broad and high jumping, kicking, throwing, and catching; also in tumbling skills such as headstands, cartwheels, forward and backward rolls, and in rhythmical skills such as skipping, hopping, sliding, and galloping to music. In University City, Missouri, a chart is posted on the gymnasium wall with each pupil's name on it (see Figure 10-1).[6] The actual

[4] Marion R. Broer, "Reliability of Certain Skill Tests for Junior High School Girls," *The Research Quarterly,* 29:139–145, May, 1958.

[5] Many of these limitations represent the professional opinion of Research Council personnel who contributed to the report, *Measurement and Evaluation Materials in Health, Physical Education, and Recreation,* American Association for Health, Physical Education, and Recreation, Washington, 1950.

[6] Jean Crowder, Helen Hartwig, and Donna Schemmer, "Charts, Public Relations and Fitness," *Journal of Health, Physical Education and Recreation,* 28:29, November, 1957.

skills are printed on the top of the chart, and as a pupil passes the skill test he colors in the appropriate square. Such a visual record provides for the teacher and pupil a quick estimate of individual progress.

For a long time there have been a number of measures of sports skills available for school and college use. Some are more reliable than others. Many are complete tests and include full instructions plus a set of practical norms. If they cannot be used directly in any one school program, they are

Names / Skills	Climb	Run	Throw	Catch	Jump	Kick	Bat	Push up	Burpee	Pull up	Cartwheel	Back roll	Head stand	High jump	Broad jump	Skip	Gallop	Forward roll	Rope jump
John																			
Bob																			
Ruth																			
Jane																			
Ellen																			
Susan																			
Rickey																			
Martin																			
Sheldon																			
Donald																			
Martha																			
Ann																			

Fig. 10-1. Chart Showing Achievement of Skill Objectives. (Courtesy Elementary Schools, University City, Mo.)

sometimes suitable for modification to meet local requirements. Once such modifications are made, however, care should be taken in the interpretation of results.

The tests that follow are representative of the better measures available. The descriptions, however, are by no means complete. This is necessary because most skill tests have details and diagrams to be adhered to rather carefully, the bulk of which are readily found in the original sources.

ARCHERY

The *Hyde Archery Test* for women was developed to evaluate achievement in archery in the Columbia Round in the 50-, 40-, and 30-yard distances.[7] It was completely evaluated on 1,400 women in twenty-seven col-

[7] Edith I. Hyde, "An Achievement Scale in Archery," *The Research Quarterly*, 8:109–115, May, 1937.

leges. Norms are available on a scale-score basis where 50 represents the average performance. These are reproduced in Table A-10, Appendix A.

In the Columbia Round a total of twenty-four arrows are shot at each distance: first range, 50 yards; second range, 40 yards; and third range, 30 yards. The standard 48-inch target is used, with the center of the gold bull's-eye 4 feet above the ground. One regular period of practice is permitted (six arrows) before competition in the Round.

In using the Columbia Round for archery appraisal, the beginning student is permitted some practice at each distance. She then takes the test, and the total score is scored on the scale for the *first* Columbia Round. After considerable practice and mastery of the event, perhaps at the close of the season or semester, the subject is again tested. This time the total score is related to the scale for the *final* Columbia Round. The score at any of the three distances can be checked against the scale score to ascertain the level of the subject. Actually, the 50-40-30–yard scale can be used during the season to note individual improvement during any practice period, and weaknesses at a certain distance can be located and corrected.

Because the Hyde Archery Test is scaled on the basis of 100 points with three standard deviations above and three deviations below the mean, it is easy to set achievement limits for estimating the quality of performance. By arbitrarily dividing the scale into five segments a group can be judged as follows: scale score 0–20, very poor; 21–40, poor; 41–60, fair or average; 61–80, good; 81–100, excellent.

BADMINTON

Badminton testing has been done for some time, and much of it has been done on individual skill elements. In 1938, Campbell devised a serving test, a forehand return test, a backhand return test, and a control test.[8] Although the test was not published, the research proved helpful to other investigators. Williams, for example, analyzed the individual tests, added several others, and found that the backhand clear gave the highest reliability of .766 and a validity of .549.[9] Other studies followed.

French-Stalter Badminton Skill Tests

In 1949, French and Stalter worked on tests to measure basic badminton skills as well as playing ability.[10] The tests are for college women. Limited norms are available in the original source, but more extensive norms are

[8] Virginia M. Campbell, "Development of Achievement Tests in Badminton," unpublished master's thesis, the University of Texas, Austin, Tex., August, 1938.

[9] Glenna Rae Williams, "A Study of Badminton Skills," unpublished master's thesis, Texas State College for Women, Denton, Tex., 1945.

[10] Esther French and Evelyn Stalter, "Study of Skill Tests in Badminton for College Women," *The Research Quarterly*, 20:257–272, October, 1949.

found in a subsequent publication.[11] The following five tests make up the battery:

1. Shuttle test: the purpose of this test is to measure footwork ability. The subject runs or side-steps between the singles side boundary lines as many times as possible in a fifteen-second period. While running she must keep her body in position to return a shuttle sent to the forehand.

2. Wrist-volley test: the purpose of this test is to measure wrist power and volleying ability. A 1½-inch wide restraining line is drawn on the floor parallel to and 6 feet from the wall. The subject stands behind the restraining line and serves the shuttle against the wall with an underhand stroke. Once the bird is served, any stroke can be used to continue volleying against

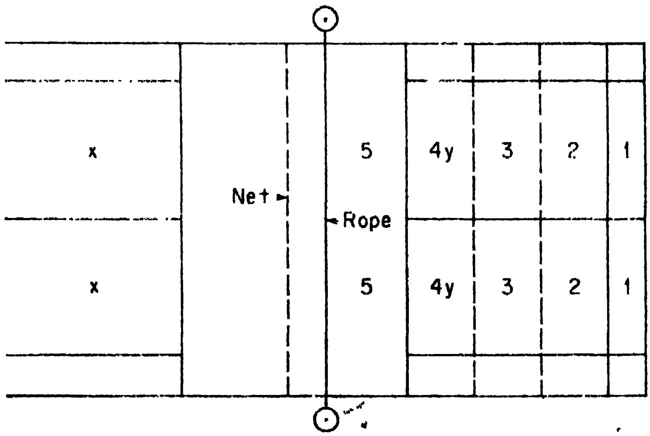


Fig. 10-2. Floor Marking for the Smash Test.

the wall. The score is kept of the number of volleys in thirty seconds. Four trials are required.

3. Smash test: in this test the ability to use the smash stroke effectively is measured. For this test a clothesline rope is stretched between two jumping standards at a height of 7 feet from the floor. The rope is drawn across the court parallel to the net and 2 feet from it (see Figure 10-2). In addition, the target is drawn on one side of a double badminton court, as indicated in Figure 10-2. The lines are painted in different colors to increase scoring accuracy. The subject to be measured (X) stands anywhere behind the short service line ready to receive twenty consecutive shuttles, which she attempts to smash between the clothesline and the net, the area marked 5 on the target. The person serving the shuttles to the subject stands at the place marked 4. The shuttle is served above the clothesline with enough force so it will readily go to the short service line opposite the target. The server acts

[11] M. Gladys Scott and Esther French, *Evaluation in Physical Education,* St. Louis, The C. V. Mosby Company, 1950, p. 69.

as scorer and calls out the score for each shuttle smashed during the twenty trials.

4. Clear test: to measure the ability to use the clear stroke effectively, the subject stands anywhere behind the short service line on the court opposite the target ready to receive twenty consecutive shuttles (see X, Figure 10-3). The target is drawn on one side of the doubles badminton court as indicated in Figure 10-3. The clothesline rope, 14 feet back from the net on the target side, is stretched between the standards 8 feet above the floor. Lines are painted different colors to improve accuracy in scoring and the target areas are numbered 4, 5, and 3. The server (Y) serves the shuttle over the net to the short service line, where the subject (X) strokes it in an effort to have it

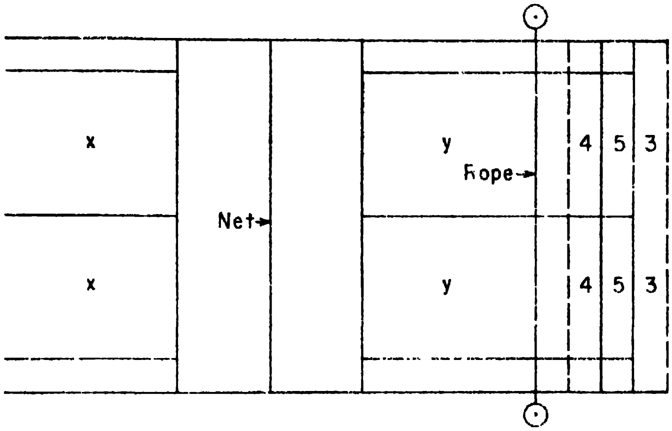


Fig. 10-3. Floor Marking for the Clear Test.

clear the clothesline and land on the target. Only good serves are played, but once a shuttle is played the trial counts. The server acts as scorer and calls out the scores for each shuttle cleared during the twenty trials. If a shuttle fails to clear the clothesline, a zero is recorded.

5. Serve test: this test measures ability to deliver the short serve accurately. The subject (X) stands anywhere in the service area diagonally opposite the target and attempts to serve the shuttle so it goes under the clothesline but over the net and lands on the target (see Figure 10-4). In this test the clothesline is strung 20 inches directly above the badminton net. The target is at the intersection of the short service line and the center line. A series of arcs 1½ inches wide are drawn in the right service court. The arc distances from the mid-point measure 22 inches, 30 inches, 38 inches, and 46 inches, respectively. The lines are painted different colors, and the target area is numbered from the corner 5, 4, 3, 2, and 1. The scorer (Y) stands nearby where she is able to observe and score each of the twenty shuttles served in the test.

A score for judging all-round playing ability is calculated on the basis of either a three-test or a four-test battery.

Three-test battery:

Playing ability score = shuttle + 0.3 (clear) + 0.19 (wrist volley)

Four-test battery:

Playing ability score = shuttle + 0.19 (wrist volley) + 0.12 (serve) + 0.11 (clear)

With little experimentation it should be relatively easy to set up standards for men. Once the target lines are laid down the test can be attempted almost any time. Badminton Achievement Scales based on T scores are found in Table A-2, Appendix A.

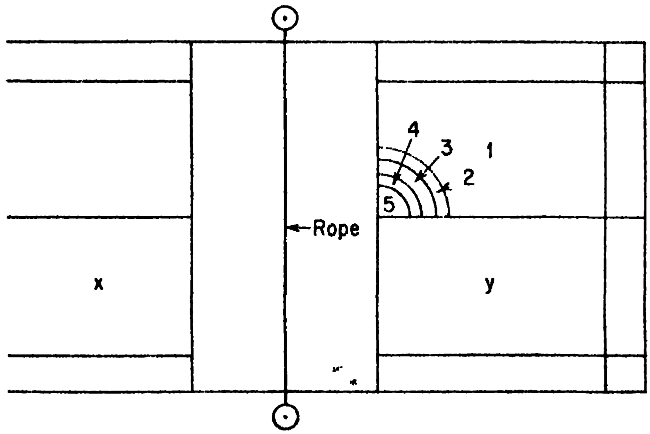


Fig. 10-4. Floor Marking for the Serve Test.

Lockhart-McPherson Badminton Test. Another test designed for college women, but which can also be used with college men, is a volley test of badminton playing ability.[12] The authors correlated their test results with the opinions of experienced badminton judges and obtained a correlation of agreement of .71. Test-retest reliability is .90.

Needed for this test is an unobstructed wall space at least 10 feet high and 10 feet long. A 1-inch line is drawn on the wall to represent the net 5 feet above the floor. A starting line is drawn parallel to the wall 6 feet 6 inches away, and a restraining line is drawn parallel to the wall 3 feet away from it. The subject to be tested stands behind the starting line with racket and shuttle, and on the "go" signal the shuttle is served against the wall and above the net line. A score is obtained by playing the shuttle against the wall as many times as possible in thirty seconds. Each player gets three trials, with rest between trials and fifteen seconds' practice before the first trial. As

[12] Aileene Lockhart and Frances A. McPherson, "The Development of a Test of Badminton Playing Ability," *The Research Quarterly*, 20:402–405, December, 1949.

the shuttle is played, the player stays within the space provided by the starting and restraining lines. Should the shuttle be missed or go out of control, the player must retrieve it and continue as before. T-score tables are available from the original source.

Miller Wall Volley Test. Rather than attempting to set up tests for a number of badminton skills, Miller analyzed the United States Amateur Badminton Championship tournament and made a careful cinematography study of all skills in an effort to locate some common ground.[13] She discovered that the finalists in both men's and women's singles events consistently used clears more often than any other stroke during all their games. Drop shots came a close second. By setting the camera up at right angles to the player, Miller was able to ascertain where the player was when the bird was hit in relation to the net.

A wall 15 feet high and 10 feet wide is needed for this test. A 1-inch line is drawn across the unobstructed wall at a height of 7 feet 6 inches. Also, a straight line is drawn on the floor 10 feet from the wall and parallel to it. After the subject has practiced for a minute he is ready to begin. On the "go" signal the subject serves the shuttlecock against the wall from behind the 10-foot line. He continues to play the bird with a clear shot on each rebound. The cumulative number of hits that land above the wall line within thirty seconds is the score. Any stroke can be used to keep the bird in play. The score consists of the sum of three trials. A short rest period of at least thirty seconds is allowed between trials.

This is a reliable and valid test, .94 and .83, respectively. It is not time-consuming, is easy to give, and can be used as a part of instruction. Moreover it works well with men. Norms need to be developed in keeping with the sexes and wall conditions.

BASEBALL

A number of factors contribute directly or indirectly to proficient baseball performance. The problem in measurement is to isolate these factors and set them up in a test battery so the type of person likely to succeed or fail in a game of baseball can be discovered. This is usually more difficult than it appears. Such factors as hitting, running, throwing, and fielding are usually observed at the secondary school level. In selecting a team from a large field of candidates, however, it might prove time-consuming to test in each of these areas. Hooks, in searching for other variables to appraise, discovered a positive relationship between body structure, muscular strength, and ability to succeed in baseball.[14] In fact, he found that certain strength meas-

[13] Frances A. Miller, "A Badminton Wall Volley Test," *The Research Quarterly*, 22:208–213, May, 1951.

[14] G. Eugenè Hooks, "Prediction of Baseball Ability through an Analysis of Measures of Strength and Structure," *The Research Quarterly*, 30:38–43, March, 1959.

ures yielded consistently high correlations with skill measures: .79, left shoulder flexion with hitting; .72, right shoulder flexion with throwing; .67, left shoulder flexion with total ability in hitting, throwing, running, and fielding. Hooks used the tensiometer to measure strength and three baseball experts to judge baseball skill performance. This is an interesting way to measure baseball ability and warrants further study.

Kelson Baseball Classification Test. This test is based on the performance of sixty-four Little League baseball players, aged eight to twelve years.[15] The measures of baseball ability selected were batting ability, throw for distance, throw for accuracy, ability to catch fly balls, and ability to catch ground balls.

Batting ability was noted by checking the batting averages for the season. In the throw for distance three trials were permitted, and the best one was selected and recorded in feet. The subject was permitted to run a little prior to throwing, but he could not step over a restraining line.

Throwing for accuracy, catching fly balls, and fielding ground balls were accomplished under the eyes of twelve judges who evaluated the performance on a 1-to-5 basis.

The outcome of Kelson's experimentation was a correlation of .85 between the baseball throw for distance and the total of five measures. In short, it appears that a carefully controlled baseball throw for distance is satisfactory for classification purposes. It has possibilities, especially for intramural tournament play, where equality of teams is recommended.

BASKETBALL

The variety of basketball tests is great. In a few years of teaching, practically every physical education teacher experiments with isolated basketball skills and their relationship to general playing ability. This has resulted in a number of "homemade" tests, many of which have been useful in screening better players for the team.

Tests for Girls

There are satisfactory tests designed to measure basketball skill in girls. One of the earliest consisted of a wall-bouncing speed test, an accuracy throw at a moving target, a free jump, a ball-handling test, and bounce-and-shoot test.[16] Later Schwartz, working with Glassow and Colvin, constructed a five-test battery that proved to be reliable and had a satisfactory validity coefficient for high school girls.[17]

[15] Robert E. Kelson, "Baseball Classification Plan for Boys," *The Research Quarterly*, 24:304–309, October, 1953.

[16] Genevieve Young and Helen Moser, "A Short Battery of Tests to Measure Playing Ability in Women's Basketball," *The Research Quarterly*, 5:3–8, May, 1934.

[17] Ruth B. Glassow and Marion R. Broer, *Measuring Achievement in Physical Education*, Philadelphia, W. B. Saunders Company, 1938, p. 113.

Leilick Basketball Test for Women. After factor analysis experimentation with the elements of basketball playing ability for college women, Leilick set up a battery of three tests: (1) bounce and shoot, (2) half-minute shooting, and (3) push pass.[18] Achievement scales were later set up by Miller and her committee.[19] These are in the form of T scores and percentile ranks based upon data collected from fifty-nine colleges. The three achievement tests are as follows:

1. Bounce and shoot: for this test two 18-foot dotted lines are drawn on the floor in the form of a V, with the point at the center of the end line under the basket. This forms a 45-degree angle. At the very end of each dotted line and perpendicular to it, a 24-inch line is drawn on the floor. Exactly 1 foot behind and 30 inches to the outside of the 18-foot line, an 18-inch line is drawn; on each of these lines is placed a chair with a basketball. Behind each chair stands a person designated as a ball catcher, whose job it is to replace the ball on the chair after receiving a pass from the person being tested.

The test begins when the subject stands at the end of the V on the 24-inch line at the right side of the basket. She picks up the ball from the chair, bounces it, shoots for the basket, recovers the rebound, and passes it back to the ball catcher from whence it came. She then runs across to the left side, picks up the ball from the other chair, and repeats the process, finally passing the rebounded ball back to the ball catcher on the left side. This operation is repeated five times on each side.

In scoring, both time and accuracy are considered. The timer notes the time to the nearest tenth of a second from the "go" signal until the subject has caught the ball after the tenth shot at the basket. The scorer measures accuracy by awarding two points when a basket is made and a single point for hitting the rim but missing the basket. Fouls, such as running with the ball, double bounce, or improper start from the 24-inch line add a penalty of one second to the time score. The final score is the best out of three trials.

2. Half-minute shooting: in a 30-second period the subject, standing at any position she chooses beneath the basket, shoots continuously as many baskets as possible. Two trials are given with two minutes between for rest. The score is the total number of baskets made in the best trial.

3. Push pass: in this event the subject, with basketball in hand, stands behind a 10-foot restraining line drawn on the floor 10 feet from a flat wall target. On the "go" signal she performs the two-hand chest pass to the target on the wall. She recovers the rebound either in the air or from the floor and continues shooting as fast as possible for a full thirty seconds. Scor-

[18] Avis Leilick, "The Primary Components of Selected Basketball Tests for College Women," unpublished doctoral dissertation, Indiana University, 1952.

[19] Wilma K. Miller, "Achievement Levels in Basketball Skills for Women Physical Education Majors," *The Research Quarterly,* 25:450–455, December, 1954.

ing is figured according to the number of target hits, weighted according to the value of the target circles. The center of the target is 4 feet 6 inches above the floor. It is a 10-inch-radius circle worth 5 points. The next concentric circle has a 20-inch radius and is worth 3 points. The outside concentric circle has a 30-inch radius with a value of 1 point. All shots must be taken from behind the restraining line. The best of two trials is recorded as the score.

Tests for Boys

If a basketball test is to find wide usage, it must be economical to give. Most tests, therefore, consist of three, four, or five skill items that can easily be given in a physical education class period.

Stroup Basketball Test. Stroup studied the basketball game scores of competing teams in 31 ten-minute games and compared them with the skill score averages of the team members.[20] The skill score averages were computed from a three-item test: passing, dribbling, and shooting. It was discovered that the basketball game is an excellent criterion for validating a particular skill test. In the Stroup study approximately 84 per cent of the games were won by the team with the higher skill score average.

Later, Stroup studied motion perception and related it to basketball playing ability.[21] This time he used a five-item skill test. He was able to demonstrate that there is a real difference between the range of motion perception of basketball players and nonplayers and that this range is a factor contributing to basketball ability.

Stroup's original three-item test, for which scale score norms are available, is as follows:

1. Foul shooting: the subject stands as near as he wishes to the basket and shoots as many baskets as possible in one minute, retrieving the ball each time himself. There are no penalties, and the score is the number of baskets made.

2. Wall passing: the subject stands behind a line 6 feet from a flat wall and passes the ball against the wall as many times as possible in one minute. The score is the number of legal passes made in one minute. It is illegal to bat the ball instead of catch it or to move ahead of the restraining line while handling the ball.

3. Dribbling: in this test the subject dribbles the ball alternately to the left and right of bottles placed in line and 15 feet apart on a 90-foot court. As he reaches the end bottle, he circles it and continues in this manner for one minute. The score is the number of bottles properly passed in one

[20] Francis Stroup, "Game Results as a Criterion for Validating Basketball Skill Test," *The Research Quarterly*, 26:353–357, October, 1955.

[21] Francis Stroup, "Relationship between Measurements of Field of Motion Perception and Basketball Ability in College Men," *The Research Quarterly*, 28:72–76, March, 1957.

minute. It is improper to knock over a bottle or to fail to pass a bottle on the proper side.

By consulting Table A-25, Appendix A, a scale score can be found for each raw score item. By averaging the three scale score items a total basketball skill score is obtained.

Lehsten Basketball Test. Working from an original group of eight tests Lehsten reduced the high school battery to five tests as follows: dodging run, 40-yard dash, baskets per minute, wall bounce, and the vertical jump.[22] Actually only two of these tests involve handling a basketball: the wall bounce and the baskets shot per minute. The wall bounce is done against a target similar to the arc in the Leilick Basketball Test for Women. The basket-shooting test is similar to the one set up by Stroup.

Johnson Basketball Test. Another test for high school boys was constructed by Johnson to measure current ability in basketball.[23] It is composed of a

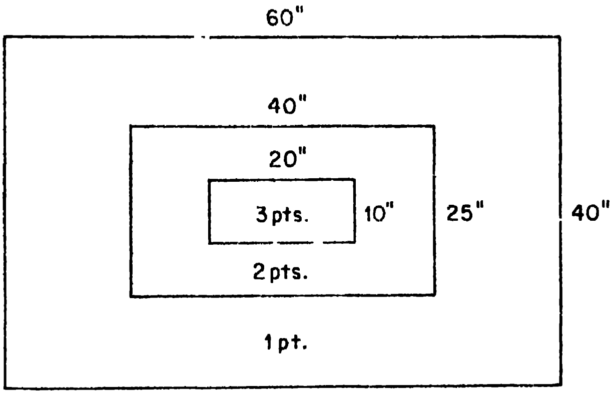


Fig. 10-5. Johnson Target for Basketball Throw.

dribble test, a throw for accuracy, and a field goal speed test. The tests were given to two groups: (1) the "good" group of basketball squad members and (2) the "poor" group, those that failed to make the squad. Employing biserial correlation Johnson obtained a validity coefficient of .88. Reliability figured to be .89.

In the *field goal test* the subject takes any position he chooses under the basket. On signal he does as many lay-up shots as possible for a period of thirty seconds. He scores a point for each basket made.

In the *throw for accuracy* the subject stands behind a restraining line 40 feet back from a wall target. The target, marked on a flat wall, contains three rectangles, one inside the other as follows: 60 by 40 inches; 40 by 25 inches; and 20 by 10 inches. The bottom of the largest rectangle is just 14 inches from the floor (see Figure 10-5). Using a basketball-type pass or hook

[22] Nelson Lehsten, "A Measure of Basketball Skills in High School Boys," *The Physical Educator*, 5:103–105, December, 1948.

[23] L. William Johnson, "Objective Test in Basketball for High School Boys," unpublished master's thesis, State University of Iowa, Iowa City, Iowa, 1934.

hold the subject throws the ball at the target. The total score resulting from ten throws is the score for this test. Points are scored as follows: 3 for the inner rectangle; 2 for the middle; 1 for the outer.

In the *dribble test* four chairs or hurdles are placed in a straight line 6 feet apart, with a distance of 12 feet from the starting line to the first hurdle. From the starting line, which is 6 feet long, the subject dribbles the basketball through the prescribed route as quickly as possible for thirty seconds (see Figure 10-6). The score is the number of hurdles passed in the period of thirty seconds.

Knox Basketball Test. One of the better tests of basketball ability was validated on high school boys who became members of the ten-man squad brought to the Oregon State Tournament by each of eight league schools.[24] The results of the experimentation speak highly for the Knox test. This is clearly demonstrated by the fact that the ten best total scores in each school were made by the ten boys who were players, and the five best total scores were made by players on the first team. Moreover, Knox was able to predict

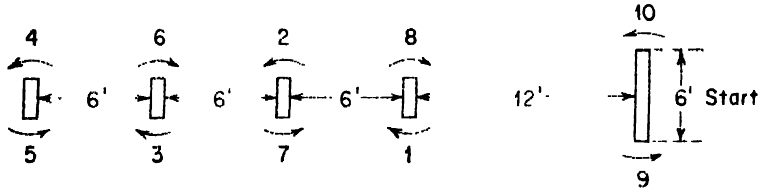


Fig. 10-6. Johnson Dribble Test for Basketball.

sixty-one out of sixty-eight squad members and twenty-nine out of thirty-six first-team members. Also, there was 89 per cent agreement between squad membership for the tournament and the scores on the basketball test.

There are four test items in the Knox test: a speed dribble, a speed pass, a dribble shoot, and a penny-cup test. A total test score is obtained by adding together the results of these four tests.

1. Speed dribble: a ball is placed on the start-finish line. The subject stands back of it with hands on knees. On "go" he picks up the ball and dribbles it down and back through a line of four chairs following a prescribed path (see Figure 10-7). The total number of seconds from start to finish is the score.

2. Dribble shoot: this test is similar to the preceding one, except that only three obstacles are used and the subject must shoot a basket before he returns. Should he fail to make a basket immediately he must keep on shooting until he is successful. The number of seconds used to complete the test is the score.

3. Speed pass: this is a rebound test done on a wall as rapidly as possible fifteen times. A line is drawn on the floor 5 feet from the wall, and the sub-

[24] Robert D. Knox, "Basketball Ability Tests," *Scholastic Coach,* 17:45–48, March,

ject employs the chest pass. The test is repeated if any rebound forces the subject to take more than one step for recovery. The number of seconds used to complete the fifteen passes is the score.

4. Penny-cup test: this reaction test is performed over a 20-foot course. There is a starting line, signal line, and finish line. The signal line is drawn on the floor 8 feet ahead of the starting line. Twelve feet further ahead is the finish line on which are placed three tin cups, painted blue, white, and red, respectively. The cups are 5 feet apart. The subject stands behind the starting line with his back to the cups and with a penny in his hand. On the "go" signal he pivots about and starts to run. As the subject crosses the signal line the tester calls a cup color, at which time the subject drops the penny into the proper cup. The test is given four times; the total elapsed time in seconds is the score.

The validity of the Knox test has been demonstrated on a number of occasions in Oregon and California, where it has been used to equate teams for basketball competition. Here the four test scores are simply added to-

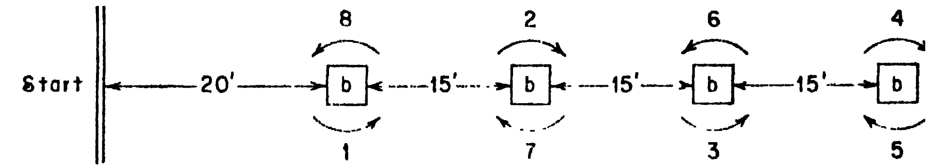


Fig. 10-7. Knox Speed Dribble Test.

gether with the low scores representing the better scores. There is little doubt that those in the low-score bracket are better players during a real game. This was shown at the University of Florida, where a biserial correlation of .96 was derived between the test scores and those men who made the squad and those who were cut from the squad.[25]

BOWLING

Bowling is a sport that lends itself very well to appraisal techniques. Controls are present for the distances, weight of the balls, and procedure. In addition each tally sheet, when kept and referred to from time to time, is a means of depicting personal improvement. About all that needs to be added is a set of norms so that comparisons can be made.

Phillips and Summers Bowling Norms. Working with 3,634 women students in twenty-two colleges, Phillips and Summers derived bowling norms for eight levels of bowling ability.[26] The average score of the first five

[25] Clifford A. Boyd, James R. McCachren, and I. F. Waglow, "Predictive Ability of a Selected Basketball Test," *The Research Quarterly,* 26:364–369, October, 1955.

[26] Marjorie Phillips and Dean Summers, "Bowling Norms and Learning Curves for College Women," *The Research Quarterly,* 21:377–385, December, 1950.

lines was selected as the criterion of ability. Ten-point intervals were established as ability levels; the lowest ability level is 50–59.9 and the highest is 120–129.9. Qualitative ratings of inferior, poor average, good, and superior were also derived based on distribution of scores through six standard deviations. The norms work well where students keep a daily record of their bowling scores.

FIELD HOCKEY

Very little has been done to measure achievement in field hockey skills. The fundamental skills, however, can be isolated and studied. Many local tests are built around one or two of the primary skills. Just prior to World War II the Schmithals-French Field Hockey Test was developed. This is a test for college women, which is divided into three sections: (1) the dribble, dodge, circular tackle, and drive; (2) goal shooting; and (3) fielding and drive.[27] The field measurements for these tests are precise and elaborate. The detailed instructions provide the girl with a chance to dribble the ball, dodge around an obstacle, execute a movement similar to a circular tackle, and finally drive the ball as hard as she can. In goal shooting the ability to shoot hard and accurately is measured. In the fielding and drive test the capacity of the girl to field a ball accurately and to execute a hard, fast drive is measured. Complete norms are not given with the original source, but norms for goal shooting and fielding drive tests are available from Scott and French.[28]

FOOTBALL

One of the chief virtues of tests is their use for standards in the physical education instructional program. With this in mind the New York State Education Department set up a number of standards in football for boys in grades seven to twelve.[29] A number of tests were established on a pass or fail basis that could be used for screening pupils in need of additional instruction. These tests are (1) receiving a forward pass (stationary) at 60 feet; (2) passing to a stationary target 60 feet away; (3) passing to a moving target moving at three-quarter speed 40 feet away; (4) passing for distance (forward direction only); (5) a center snap of 24 feet that reaches the receiver between the shoulders and knees; (6) a punt for distance (in feet); (7) a placement kick through goal posts at 60-foot distance; and (8) a kickoff for distance (in feet). The eight tests are not difficult to administer and have had considerable use in New York State public schools.

[27] Margaret Schmithals and Esther French, "Achievement Tests in Field Hockey for College Women," *The Research Quarterly*, 11:84–92, October, 1940.

[28] M. Gladys Scott and Esther French, *op. cit.*, p. 91.

[29] New York State Physical Education Standards Project, *Standards Manual for Football, Soccer, and Softball Skills, Boys, Grades 7–12,* Bulletin 3, New York State Education Department, Albany, N.Y., 1951.

Borleske Touch Football Test. Touch football playing ability was studied by Borleske, who sampled the opinion of forty-six judges on the important skills in the sport.[30] As a result of their combined judgments five test items were selected, which produced a correlation of .925 with the criterion: (1) forward pass for distance; (2) catching a forward pass; (3) punting for distance; (4) 50-yard run carrying a ball; and (5) pass defense.

Later the five items were reduced to three with only a small loss in test validity (.85). These are the tests.

1. Forward pass for distance: on a field marked with lines every 5 yards and with markers every 10 yards, two subjects can throw the football from both ends of the field, thus saving time. Each subject checks the place where the partner's ball hit the field and estimates the distance to the nearest yard. Three throws are permitted after a one-minute warm-up. Each throw starts by the catch of a pass from center. The best of three throws is scored.

2. Punt for distance: punters work together at either end of the field similar to the previous test. The punt is preceded by a pass from center; within two seconds after the ball is received, it must be kicked. The best of three trials is scored.

3. Fifty-yard speed sprint: the subject assumes a backfield three-point stance 5 yards back of the center. He receives the ball on a snap from center and carries it by any form as he sprints for a distance of 50 yards. A minute is permitted for warm-up activity. Time is recorded in seconds. T-score scales are available.[31]

Brace Football Achievement Test. Brace, working with college men, designed an eight-item skill test for preselecting players most likely to succeed in football.[32] The items are a forward pass at a target, a 50-yard sprint carrying a football, a forward pass for distance, a pull-out, blocking, a punt for distance, a dodge and run, and charging. Instructions for administering the test plus the extensive field layout are quite lengthy and will be omitted here. The items, however, are suitable for both college and high school boys. Norm tables are not available.

GOLF

Despite the great amount of golf instruction given in the country, there are not many standardized tests. Perhaps one reason for this is that a properly completed golf score card, representing any one person on any one golf

[30] Frederick W. Cozens, "Ninth Annual Report of Committee on Curriculum Research of the College Physical Education Association: Part III," *The Research Quarterly*, 8:73–78, May, 1937.

[31] Achievement scales are listed in the cited source or may be found in report by S. E. Borleske, *A Study of Achievement of College Men in Touch Football*, Berkeley, Calif., University of California Press, 1936.

[32] David K. Brace, "Validity of Football Achievement Tests as Measures of Motor Learning and as a Partial Basis for the Selecting of Players," *The Research Quarterly*, 14:372–377, December, 1943.

course, is in itself an appraisal instrument, particularly if it is not discarded at the end of the match. Moreover, it is relatively easy to put together a local test of putting, driving for distance, and making chip shots to the playing green.

It is difficult, even with stroboscope pictures, to analyze the golf drive completely. The factor of individual differences shows up in the movement of professional golfers. Rehling found this to be true in his study of 2,500 pictures and twelve golf professionals in Florida.[33] It seems, therefore, that no one method of hitting the golf drive can be called the correct method.

Homemade targets, patterned after an archery target, hanging in a caged-in area have been used for years to stimulate player accuracy in shooting or driving. Also, for the right-handed player, shots that strike the left-hand side of the target lean toward the hooked shot, and those that strike the far right of the target are considered sliced shots. For putting practice, collapsible floor cups can be set about a smooth gymnasium floor or front lawn of the local school building. A simple scoring technique to give credit for the more difficult shots is all that is necessary to keep interest high in golf classes.

Clevett did some experimenting with indoor tests back in 1931, in which he used a 10-foot square target, marked off with twenty-five numbered areas and placed in a cage or net some twenty-one feet from the tee-off spot.[34] A midiron is used by the subject. Ten trials are taken right away without practice or warm-up. A similar test is performed ten times using a mashie to simulate the short approach shot to the green. Finally there is a 20-foot long putting test done on a 27-inch-wide piece of floor carpet divided into forty-eight separate scoring areas. However, it would not be hard to improve on this test by using the collapsible golf cup in the modern gymnasium, where numbers could be placed on the floor about the cup. Clevett did not construct norms, but in any of these measures the subject can judge his performance against previous performances.

GYMNASTICS

Ability in gymnastics is very nicely judged during interscholastic or intercollegiate competition. When three or more experienced judges rate a performer on his ability in a certain event, there is a satisfactory degree of control. The judgment is usually far less subjective than it seems to be from the side lines of the gymnasium. In most progressive gymnastic programs a series of prescribed skills are posted near the various pieces of equipment, and a student can be checked at almost any time on his proficiency.

[33] Conrad H. Rehling, "Analysis of Techniques of the Golf Drive," *The Research Quarterly,* 26:80–81, March, 1955.
[34] Melvin A. Clevett, "An Experiment in Teaching Methods of Golf," *The Research Quarterly,* 2:104–106, December, 1931.

Form in the prescribed gymnastic events is what makes better performers. Zwarg, a master performer and instructor in this specialty for many years, has done much to improve the objectivity of gymnastic ratings.[35] Competent judges today seldom show poor objectivity as they rate performers. Hunsicker and Loken demonstrated the truth of this statement at an N.C.A.A. meet.[36]

Wettstone, of Pennsylvania State University, seems to be one of the few persons to have experimented with ways of predicting gymnastic and tumbling ability.[37] He found that the good gymnast, as judged by gymnastic coaches, had a high level of motor capacity as demonstrated by thigh flexion strength, chin-ups, dips, squat-thrusts, and thigh circumference divided by height. This is not difficult to accept, for in any cross section of young men, those rating high in physical fitness items, such as those used by Wettstone, might be expected to do very well in gymnastics. A number of other athletes, however, should do just as well. In such case, and even if they had never been on a piece of gymnasium equipment, they would show a potential strength for gymnastics.

HANDBALL

Handball is one of those sports, like tennis, squash rackets, and badminton, that lend themselves to round-robin and doubles tournament play. Individual achievement in handball can be checked by where the subject stands when the tournament is finished. This, of course, does little for instructional handball.

Montoye and Brotzman studied the tournament means of measuring skill in handball and came to the conclusion that a doubles handball tournament with partners selected at random, no two players forming a team more than once, ranks the players according to ability in approximately the same order as a singles tournament, providing the number of games is sufficiently large and that either the average score or the difference between player's score and that of his opponents is used.[38] Here, each subject should play eleven or twelve matches for increased validity.

Cornish Handball Test. After studying five handball test items Cornish narrowed his test down to two events: (1) power test and (2) thirty-second

[35] Leopold F. Zwarg, "Judging and Evaluation of Competitive Apparatus for Gymnastic Exercises," *Journal of Health, Physical Education and Recreation,* 6:23–25, January, 1935.

[36] Paul Hunsicker and Newt Loken, "The Objectivity of Judging at the National Collegiate Athletic Association Gymnastic Meet," *The Research Quarterly,* 22:423–424, December, 1951.

[37] Eugene Wettstone, "Tests for Predicting Potential Ability in Gymnastics and Tumbling," *The Research Quarterly,* 9:115–118, December, 1958.

[38] Henry J. Montoye and John Brotzman, "An Investigation of the Validity of Using the Results of a Double Tournament as a Measure of Handball Ability," *The Research Quarterly,* 22:214–218, May, 1951.

volley test.[39] These two events alone correlated with the criterion measures with a correlation of .667. In the power test the subject is awarded points for serving a ball with enough power to reach the back court. In the thirty-second volley test the object is to volley the ball against the front wall and back to the service line as often as possible in the time alloted. Evidence on the use of this test is missing, and norm tables are not presently available.

RIDING

The equestrian art has always had its judges. From the age of chivalry to the present county horse show, riders have been appraised for form and riding ability. In fact, many a camp or college instructor evaluates his charges on the basis of horsemanship, an all-round term covering general knowledge and skill both on and off the animal.

As in diving and gymnastics, form is extremely important. This makes it difficult to set up a riding test and probably accounts for the lack of satisfactory appraisal techniques.

Crabtree Riding Test. Crabtree, on an empirical basis, developed a nine-item score chart that can be used to judge riding in competition.[40] A maximum of 40 individual points can be accumulated. These are distributed as follows:

Mounting (3 points)	Reins (1½), ease and coordination (1½)
Dismounting (3 points)	Reins (1½), ease and coordination (1½)
Walk (5 points)	Hands (1), circle (1), reverse (1), form (2)
Trot (14 points):	
Collected trot	Change diagonals (2), figure eight (2), circle (1), reverse (1), elbows (1), hands (1), legs (1), heels (1), rhythm (1)
Extended trot	Seat (1), hands (1), control (1)
Canter (11 points):	
Collected canter	Right lead (1), left lead (1), figure eight (2), seat (2), hands (2)
Extended canter	Seat (1), hands (1), control (1)
General horsemanship (4 points)	A general opinion item

In the Crabtree test the subject is judged on performance and rated on the scale. In competition this appears to be a fair means of rating all types of riders at the same time. As a teaching aid the score sheet has value to the

[39] Clayton Cornish, "A Study of Measurement of Ability in Handball," *The Research Quarterly*, 20:215–222, May, 1949.

[40] Helen Kitner Crabtree, "An Objective Test for Riding," *Journal of Health and Physical Education*, 15:419, October, 1943.

student, for a judge or instructor can offer concrete criticism to each person before they leave the riding ring.

RHYTHM TESTS

Rhythm appraisal in its early stages is as much the domain of the music teacher as it is of the physical education teacher. Both specialists are concerned with the motor response to sound in its many and varied forms. Rhythm in the more advanced stages is embodied in the patterns of creative, interpretive, ballet, folk, social, and contemporary dance.

Although individual rhythmic ability can be noted in the classroom or on the playground, it is not easy to measure objectively. It appears to be an individual, kinesthetic, and instinctive experience. It does, however, contribute to a child's advancement at the preschool level.[41] It appears to be related to the speed of reaction time.[42] In fact, motor response has generally been associated with rhythmic ability. The work of Benton illustrates this conclusively.[43] Ashton also demonstrates this. He showed that gross motor rhythm ability could be measured by having students walk, run, and skip to music, by making improvisations on musical excerpts, and by combining dance steps, where the subject initiates the movement for waltz, schottische, or polka after the step is verbally identified.[44]

Waglow constructed a social dance test that was validated on the performances of 60 men and women who had just completed the social dance course; they were scored by two independent scorers.[45] The coefficient of validity was only .376, and the reliability coefficient was .472. Further investigation into the scoring system and the consistency of performances is necessary. In this test the subject listens to a few measures of a dance rhythm and then dances either a waltz, tango, slow fox trot, jitterbug fox trot, rumba, or samba. In all tests the objectivity was highest in the waltz.

Simpson developed a Rhythmeter to measure locomotor response to auditory rhythmic stimuli.[46] Using dance as a locomotor activity women were

[41] Studies from the Center for Research in Child Health and Development, School of Public Health, Harvard University, *Monograph of the Society for Research in Child Development,* vol. IV, serial no. 21, National Research Council, Washington, 1939.

[42] Don J. Wilson, "Quickness of Reaction and Movement Related to Rhythmicity or Nonrhythmicity of Signal Presentation," *The Research Quarterly,* 30:101–109, March, 1959.

[43] Rachel Jane Benton, "The Measurement of Capacities for Learning Dance Movement Techniques," *The Research Quarterly,* 15:137–142, May, 1944.

[44] Dudley Ashton, "A Gross Motor Rhythm Test," *The Research Quarterly,* 24:253–258, October, 1953.

[45] I. F. Waglow, "An Experiment in Social Dance Testing," *The Research Quarterly,* 24:97–101, March, 1953.

[46] Shirley E. Simpson, "Development and Validation of an Objective Measure of Locomotor Response to Auditory Rhythmic Stimuli," *The Research Quarterly,* 29:342–348, October, 1958.

appraised for rhythmic ability. Performance scores were consistent. Individual subject movements in response to fifteen rhythmic patterns are transferred to a kymograph drum, where an easy comparison is made with the master mark. A response within 1 millimeter of the master mark is considered correct. The music for the dance steps was played on the piano and recorded on a magnetic tape. With more experimentation the instrument may fulfill a need in this area.

SOFTBALL

For instructional purposes softball skills can be appraised in about six basic ways: (1) batting a pitched ball thrown from 40 feet away; (2) throwing for accuracy up to 60 feet; (3) catching a line ball up to 60 feet away from the thrower; (4) catching a ground ball up to 60 feet away from the thrower; (5) catching a fly ball up to 140 feet away; and (6) throwing for distance. These tests are from the New York State Physical Education Standards Project and are used in the secondary schools of New York State.[47] Similar items appear in the older Rodgers-Heath Softball Test, which is limited to fifth and sixth grade boys.[48]

O'Donnell Softball Skill Test. This is a measure for high school girls developed by O'Donnell for softball classification purposes.[49] Six tests are employed.

1. Speed throw: The subject stands behind a 5-foot-long restraining line. On the "go" signal she reaches down, picks up a ball, and throws it as hard and fast as she can toward a flat wall 65 feet away. She may not step over the line until the ball is actually released. The score is the time in seconds it takes the ball to hit the wall. Three trials are given.

2. Throw and catch: A rectangle, 20 by 50 feet, is necessary for this test. The rectangle is divided the long way into ten equal areas by lines 5 feet apart. One end of the rectangle is the starting end. Across this end is strung a piece of rope 8 feet high and parallel to the floor. The subject stands behind the rope and throws the ball over the rope and runs to catch it on the other side. The attempt is to cover as much ground as possible. Any throw is permitted, and three trials are given. The score is the distance in feet from the starting line to the point at which the ball is caught.

3. Repeated throws: For this test a line 7 feet 6 inches above the floor and 15 feet long is drawn on a flat wall. On the "go" signal the subject, who stands behind the 15-foot line, picks up the softball and throws it continuously against the wall. The attempt is to hit the wall above the wall line as

[47] New York State Physical Education Standards Project, *op. cit.*, pp. 23–25.

[48] Elizabeth G. Rodgers and Marjorie L. Heath, "An Experiment in the Use of Knowledge and Skill Tests in Playground Baseball," *The Research Quarterly*, 2:113-115, December, 1931.

[49] Doris J. O'Donnell, "Validation of Softball Skill Tests for High School Girls," unpublished master's thesis, Indiana University, Bloomington, Ind., August, 1950.

often as possible in a thirty-second testing period. One trial is given, and the number of hits is the score.

4. Fielding fly balls: This test is similar to the one above, except that a line is drawn on the wall parallel to the floor at the 12-foot level. The subject stands behind the restraining line and throws the ball continuously at the wall and over the 12-foot line. She catches the rebounding ball anywhere but must always return behind the restraining line to make the next throw. The test is scored on the balls caught during a thirty-second trial.

5. Fungo batting: The subject stands in the batter's box, ball in one hand and bat in the other. She tosses the ball into the air and tries to hit it as it comes down. She is given ten chances (trials). A swing and a miss counts as a trial. A ball hit to the outfield is scored 5 points; one to the infield, 3 points; and a foul ball, 1 point.

6. Throw for accuracy: A wall target is attached to or painted on a wall with its center 3 feet from the floor. The four-circle target is graded from the center 4, 3, 2, 1. A 5-foot restraining line is made on the floor 45 feet away from the target. From behind this line the subject makes an overhand throw toward the target and is scored accordingly. Ten throws are given to obtain a total test score. Stepping over the line forfeits the throw.

These six tests provide a reasonably good measure of softball ability. They are interesting to take and can be worked into the teaching program with ease. O'Donnell found that a three-test and two-test battery was effective in appraising skill:

Three-test battery:

Softball playing ability score = 0.6 (throw and catch)
+ 1.1 (repeated throws) + 0.5 (throw for accuracy)

Two-test battery:

Softball playing ability score = 0.8 (throw and catch)
+ 1.7 (repeated throws)

Norms were made on so few cases that it is recommended that test scores be scaled from the raw scores at the local level. The tests may have far more meaning this way in a particular community.

SOCCER

A simple soccer test at the elementary school level will select the poor players in need of additional instruction and do more for intramural soccer than almost any other single factor. The Heath and Rodgers test is aimed toward fifth and sixth grade boys; it is a combination of knowledge and skill tests.[50] The subject is asked to kick a rolling ball, place kick for the goal, demonstrate a throw-in, and execute a dribble.

[50] Marjorie L. Heath and Elizabeth G. Rodgers, "A Study of the Use of Knowledge and Skill Tests in Soccer," *The Research Quarterly*, 3:33–37, December, 1932.

In the New York State Physical Education Standards Project the subject must (1) execute a corner kick to within the goal or penalty area; (2) execute an instep kick for distance; (3) dribble the ball at jogging speed for 75 feet; (4) execute an instep kick from bounce for distance; and (5) carry out a tow kick for distance. Standards of achievement for four levels are available for boys in grades seven to twelve.[51]

McDonald Soccer Test. McDonald was interested in predicting soccer game proficiency in college players.[52] College coaches acting as judges of playing ability turned in appraisals which were correlated with performance on the prescribed soccer test. Validity ranged from .63 to .94, depending on the type of college group surveyed.

For the test a soccer backboard is used. It must be 30 feet wide and 11½ feet high. Nine feet in front of it a restraining line is drawn. The object is for the subject to kick the ball against the backboard as often as possible in a 30 second period. Any variety of kick is permitted as long as the supporting leg remains behind the restraining line. The best 30 second score in four trials is recorded. A retrieved ball has to be brought back to the restraining line before the test can be continued.

SPEEDBALL

About the only test of speedball skills that has been used to any extent is the one by Buchanan developed for use with high school girls.[53] It is divided into four parts: (1) lifting the ball accurately with either foot and passing it; (2) throwing and catching against a wall; (3) dribbling and passing while moving down the field; and (4) kick-up of a moving ball. A complete breakdown of the test, together with its diagrams and limited norms, is found in the Weiss and Phillips manual.[54]

SWIMMING

No more elaborate program of progressive skill testing exists in sport than in swimming. The very carefully developed series of swimming tests of the Young Men's Christian Association and the long-standing tests of the American National Red Cross are known to vast numbers of persons in and out of physical education. The tests are difficult to improve upon. Organizations such as the boy scouts and girl scouts have merely modified existing tests to suit their particular needs.

[51] New York State Physical Education Standards Project, *op. cit.*, pp. 55–80.

[52] Lloyd C. McDonald, "The Construction of a Kicking Skill Test as an Index of General Soccer Ability," unpublished master's thesis, Springfield College, Springfield, Mass., 1951.

[53] Ruth E. Buchanan, "A Study of Achievement Tests in Speedball for High School Girls," unpublished master's thesis, State University of Iowa, Iowa City, Iowa, 1942.

[54] Raymond A. Weiss and Marjorie Phillips, *Administration of Tests in Physical Education*, St. Louis, The C. V. Mosby Company, 1954, pp. 253–257.

Cureton, a veteran researcher in the study of tests of swimming ability, has shown that swimming performance is related to circulatory-respiratory tests and that energy metabolism tests are significantly related to 100-yard and 440-yard swim times. Physical condition is so important to swimming skill results that Van Huss and Cureton, in analyzing the 100-yard swim, found that muscular endurance is more important than the relative contribution of arm stroke and leg kick.[55]

Hewitt Swimming Achievement Scales. Seeking a test to classify students into homogeneous swimming group, Hewitt analyzed the performances of college men.[56] He built his test on the following items:

Underwater swim, 20 and 25 yards
Endurance swim
25-yard sprint, crawl, breast, and back crawl strokes
50-yard swim, crawl, breast, and back crawl strokes
50-yard glide relaxation swim, elementary back, side, and breast strokes

A correlation of validity was established at .87 between the criterion and the 25-yard (or 50-yard crawl) plus the three gliding strokes. Shortly after this test was worked out, Hewitt set up a reliable and valid test for high school boys and girls.[57] It consists of the 50-yard crawl, a 25-yard flutter kick, and a 25-yard glide relaxation for the elementary back stroke, side stroke, and breast stroke. The side stroke had the highest relation to the total test.

Fox Test of Swimming Power. This test was established on the power of the side stroke and front crawl; reliability is .97 and .95, respectively, and validity is satisfactory.[58] This test for college women consists of two items, as follows:

1. Side stroke: The subject starts in the pool in a side float position with the malleoli resting on a taut rope strung across the pool, 2 feet from one end. There is a little weight hung on the rope so when the free end is dropped it will submerge easily. When the rope is dropped the subject takes exactly five complete side strokes emphasizing the glide for distance. She is scored in feet according to the distance recorded along the side of the pool deck (which is marked off in 5-foot intervals). The marking position is the location of the ankles at the beginning of the recovery of the legs for the sixth stroke.

2. Front crawl: The start is from a prone float position with the feet at

[55] W. D. Van Huss and T. K. Cureton, "Relationship of Selected Tests with Energy Metabolism and Swimming Performance," *The Research Quarterly,* 26:205–221, May, 1955.

[56] Jack Hewitt, "Swimming Achievement Scale Scores for College Men," *The Research Quarterly,* 19:282–289, December, 1948.

[57] Jack Hewitt, "Achievement Scale Scores for High School Swimming," *The Research Quarterly,* 20:170–179, May, 1949.

[58] Margaret G. Fox, "Swimming Power Test," *The Research Quarterly,* 28:233–237, October, 1957.

malleoli level on the rope. Five complete strokes are taken. The distance covered is measured by noting where the ankles are at the moment the fingers enter the water to begin the sixth complete arm cycle. This can be simplified by counting the point of entry of each arm as one unit. The eleventh entry is the time to take the measurement.

T-scale scores are available in the original source. Additional work indicates that scales for the back crawl, breast stroke, and elementary back stroke will be forthcoming. Briefly, the scale scores break down in large blocks as follows:

T score	Side stroke, feet	Crawl, feet
75	52	50
50	28	27
25	5	8

TABLE TENNIS

A table tennis skill test was devised by Mott and Lockhart following the pattern displayed in volleying tests, where an object is volleyed as many times as possible against a perpendicular wall.[59]

One end of a hinged table is lifted and made secure so the surface is perpendicular to the other end of the table. This affords a flat wall upon which to hit a ping-pong ball. A chalk line is drawn horizontally across this surface 6 inches up the wall to represent a net. Extra balls are placed in an open kitchen matchbox and attached to the player's end of the table flush with the surface. On the "go" signal the subject volleys the ball against the perpendicular wall of the table just as rapidly as possible for thirty seconds. If she loses a ball, she may substitute another from the matchbox nearby. Her free hand must not lean on the table, and no ball may be returned until it takes at least one bounce. The score is the best score of three trials.

Norm tables are available. A raw score over 56 volleys in thirty seconds would place the girl in the top quartile. A score of 39–55 volleys represents the third quartile; 20–38, the second quartile; and below 20, the bottom quartile.

TENNIS

Experimentation in tennis testing has been going on for several decades. Few tests in any skill enjoy the well-earned prestige of the Dyer Tennis Test.

[59] Jane A. Mott and Aileene Lockhart, "Table Tennis Backboard Test," *Journal of Health and Physical Education,* 17:550, November, 1946.

Dyer Tennis Test. This test is simply a backboard volleying measure, the score of which is significantly related to over-all tennis ability.[60] For years the test has been used to rate varying degrees of tennis ability for men and women of most ages. Dyer obtained a correlation of .92 between test scores and status in round-robin tournament play. A .53 correlation between beginning college women players' test scores and subjective ratings of several tennis skills was obtained by Fox.[61] This correlation coefficient was increased to .81 when the Dyer test was combined with the Miller Forehand Backhand Test. Shay vouches for the worth of the Dyer test by reporting that, when careful records were kept at Williston Academy, the six boys with the highest test scores in the institution became the first six players on the academy team and were never displaced.[62] Moreover, in a group of sixty boys classified into three groups according to test scores only two members from the lower groups were ever able to advance to the first-team level.

In planning space for this test a 3-inch wide line is drawn on the backboard 3 feet above the floor to represent the net. A line 15 feet in length is drawn 5 feet away from the backboard and parallel to it.

The subject to be tested stands behind the restraining line with two balls in the hand. On the "go" signal a ball is dropped to hit the floor once and then put in play against the wall. The object is to hit the ball as many times as possible in thirty seconds. Each ball to score must strike the wall on or above the net line. It may be volleyed directly, or it may take several bounces. Any stroke may be employed as long as the subject keeps behind the restraining line. Any number of balls may be used. The final test score is the sum of the scores on three trials. This can be related to a scoring table to give it a more specific meaning (see Table A-7, Appendix A).

Broer-Miller Tennis Test. By concentrating on backhand and forehand driving ability, a test was developed to classify women students in tennis.[63]

In planning space for this test a regulation court with net is used. Running parallel to the net and 4 feet above it is a piece of rope. A chalk line is drawn across one end of the court 10 feet inside the service line and parallel to it (see Figure 10-8). Beyond the base line two more parallel lines are drawn 5 feet apart. Numbers are placed in the center of each area to designate the scoring value.

The subject to be tested stands behind the base line with a good ball. When ready, she bounces the ball, hits it, and tries to place it in the back 9

[60] Joanna T. Dyer, "The Backboard Test of Tennis Ability," *The Research Quarterly* (Supplement), 6:63-65, March, 1935; "Revision of Backboard Test of Tennis Ability," *The Research Quarterly*, 9:25–29, March, 1938.

[61] Katherine Fox, "A Study of the Validity of the Dyer Backboard Test and the Miller Forehand Backhand Test for Beginning Tennis Players," *The Research Quarterly*, 24:1–7, March, 1953.

[62] Clayton T. Shay, "An Application of the Dyer Tennis Test," *Journal of Health and Physical Education*, 20:273–274, April, 1949.

[63] Marion R. Broer and Donna Mae Miller, "Achievement Tests for Beginning and Intermediate Tennis," *The Research Quarterly*, 21:303–310, October, 1950.

feet of the opposite court. Fourteen trials are given for each style: forehand and backhand. This is accomplished by striking the ball so it passes between the top of the net and the restraining rope placed above the net. The more accurate, flat, and powerful the drive, the better the score. The sum of the scores of the twenty-eight balls is the total score.

Validity and reliability are .85 and .80, respectively. The authors feel that the test is a better test for intermediates than it is for beginners. As it does

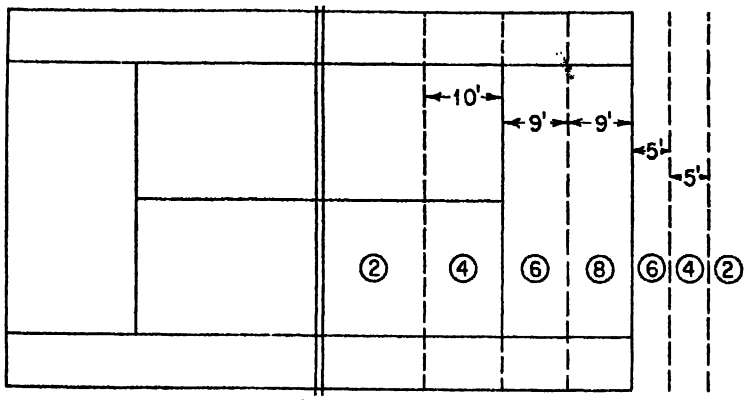


Fig. 10-8. Court Markings for Broer-Miller Tennis Test.

not correlate highly with the Dyer test, it probably is a measure of other qualities of performance. Like the Dyer test, however, it correlates well with demonstrated playing ability.

VOLLEYBALL

One of the earliest volleyball tests was developed in Wisconsin by Bassett, Glassow, and Locke.[64] Like so many of the earlier tests, it was validated on college women. It consists of a serving test to a fixed target and a volleying test against a wall for thirty seconds. Reliability and validity could be higher, but the test items are useful. French and Cooper verified the value of these two volleyball tests and made improvements on them.[65]

Russell-Lange Volleyball Test. Studying girls in grades seven, eight, and nine, Russell and Lange worked on the serving and volleying tests of earlier researchers and obtained reliability and validity coefficients of over .80.[66]

Volleying Test: For this test a 10-foot-long line is marked on the wall 7½ feet above the floor. Another line 10 feet long is drawn on the floor 3 feet

[64] Gladys Bassett, Ruth Glassow, and Mabel Locke, "Studies in Testing Volleyball Skills," *The Research Quarterly,* 8:61–72, December, 1937.

[65] Esther L. French and Bernice I. Cooper, "Achievement Tests in Volleyball for High School Girls," *The Research Quarterly,* 8:150–157, May, 1937.

[66] Naomi Russell and Elizabeth Lange, "Achievement Tests in Volleyball for High School Girls," *The Research Quarterly,* 9:33–38, December, 1940.

out from the wall. The subject to be tested stands behind this 3-foot restraining line and tosses the ball to the wall. As the ball rebounds, the subject volleys it as many times as possible against the wall above the net line in a thirty-second period. The ball may be played in any manner, but if it goes out of control, the subject must bring it back on test time. The number of times the ball is struck from behind the restraining line to the proper area on the wall is the score. The best score from three trials is recorded.

Serving Test: Special court markings in chalk are needed for this test (see Figure 10-9).

The subject to be tested stands to the rear of the base line in the serving area. Using any proper volleyball service, the subject serves a total of ten balls over the net, attempting to score the greatest number of points possible

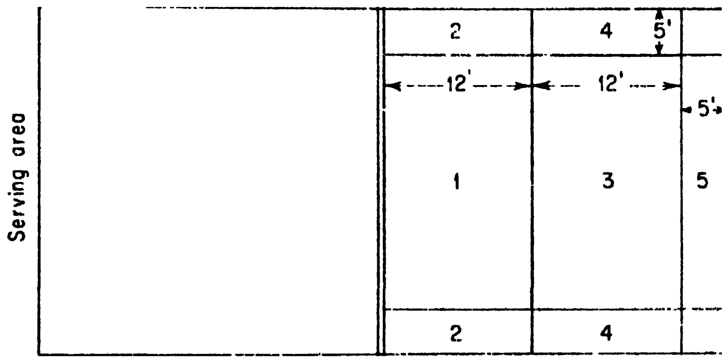


Fig. 10-9. Markings for Russell-Lange Serving Test.

with each serve. A let ball is served over. Foot faults score zero. Two trials of ten serves are given each subject. The best trial is recorded.

Volleyball playing ability is the total of the scaled scores for the two tests. Scoring tables for junior high school girls appear in Table A-22, Appendix A.

Brady Volleyball Test. An easy-to-administer volleying test was designed by Brady for college men.[67] On a smooth wall 11½ feet above the floor a 5-foot-long horizontal line is drawn. At each end of this line vertical lines are extended to the ceiling.

The subject to be tested may stand any place he chooses and throw the ball against the wall. Having done this, he volleys continuously for one minute, scoring only those hits that land within the target area. All volleys must be legal.

To use the test for grading purposes Brady takes the difference between the scores made on the firts test and the last test and adds it to the score of

[67] George F. Brady, "Preliminary Investigations of Volleyball Playing Ability," *The Research Quarterly,* 16:14–18, March, 1945.

the last test. The weak beginner who learns rapidly is thus rewarded, without in anyway penalizing the good beginner.

GENERAL SPORTS SKILLS

By combining two or more batteries of skill tests a single battery for appraising general skill ability can be constructed. Such a battery is a grouping of specific sport skills and is not to be confused with a motor ability test. Johnson, for instance, examined the test items from five team sports: volleyball, touch football, softball, soccer, and basketball.[98] He then built a general sports skill test that included the Borleske football pass for distance, the Knox basketball dribble, a Brady-type volleying test for volleyball, a softball throw for distance, and a soccer volley test against a wall. He then developed a regression equation from which to calculate a total score.

> 2.8 (basketball dribble) + .6 (soccer repeated volley) + 1.1 (softball throw for distance) + 1.0 (touch football distance pass) + .4 (volleyball volleying)

An average score would fall between 203 and 286 points, with a good score falling above these limits and a poor score falling below.

GENERAL ACHIEVEMENT SCALES

Achievement scales have usually been derived for the purpose of classifying pupils for instructional classes, selection of teams, and local research. They frequently include a wide variety of fundamental abilities as well as special sports skills.

Cozens Scales. Over a period of years Cozens and his fellow workers have brought out a wide variety of scales covering all grade levels and both sexes. Pupil classification is in keeping with the Cozens classification index. All the Cozens scales are built on a 6-sigma score organization. They include the following:

1. *Achievement Scales in Physical Education Activities for Boys and Girls in Elementary and Junior High Schools* (with N. P. Neilson, New York, A. S. Barnes and Company, 1934). There are twenty scaled events for girls, thirty-three for boys: They include soccer, basketball, baseball, track and field items, and boys' strength items.

2. *Achievement Scales in Physical Education Activities for Secondary School Girls and College Women* (with N. P. Neilson and H. J. Cubberly, New York, A. S. Barnes and Company, 1937). These include a wide variety of scales for jumping, hopping, speedball, swimming, field hockey, volleyball, archery, tennis, baseball, basketball, running, and soccer.

[98] Kenneth P. Johnson, "A Measure of General Sports Skill of College Men," unpublished doctoral dissertation, Indiana University, Bloomington, Ind., 1956.

3. *Physical Achievement Scales for Boys in Secondary Schools* (New York, A. S. Barnes and Company, 1936). Scales cover walking, track and field, muscle strength, gymnastics, football, basketball, and baseball.

4. *Achievement Scales in Physical Education Activities for College Men* (Philadelphia, Lea & Febiger, 1936). The tests are practically the same as for high school boys.

McCloy Achievement Scales. McCloy's scoring tables for boys and young men follow much the same pattern as those of Cozens. They are set up according to age level and are based on the McCloy Classification Index (see Charles H. McCloy, *The Measurement of Athletic Power,* New York, A. S. Barnes and Company, 1932).

SELECTED REFERENCES

Bovard, John F., Frederick W. Cozens, and E. Patricia Hagman: *Tests and Measurements in Physical Education*, Philadelphia, W. B. Saunders Company, 1949, pp. 312–324.

Broer, Marion R.: "Evaluation of a Basic Skills Curriculum for Women Students of Low Motor Ability at the University of Washington," *The Research Quarterly*, 26:15–27, March, 1955.

Cozens, Frederick W., and Florence A. Stumpf: *Sports in American Life*, Chicago, University of Chicago Press, 1953.

Jones, Harold E.: *Motor Performance and Growth*, Berkeley, Calif., University of California Press, 1949, pp. 160–170.

Larson, Leonard, and Rachel Yocum: *Measurement and Evaluation in Physical, Health and Recreation Education*, St. Louis, The C. V. Mosby Company, 1951, chap. 9.

Latchaw, Marjorie: "Measuring Selected Motor Skills in Fourth, Fifth, and Sixth Grades," *The Research Quarterly*, 25:439–449, December, 1954.

Martin, Joan L.: "Bowling Norms for College Men and Women," *The Research Quarterly,* 30:113–116, March, 1960.

Walters, Etta C.: "Motor Ability and Educability Factors of High and Low Scoring Beginning Bowlers," *The Research Quarterly,* 30:94–100, March, 1959.

CHAPTER 11

General Motor Ability and Motor Intelligence

General motor ability is a fascinating topic, for sooner or later in the study of it one is bound to cross trails with most educational disciplines and numerous characteristics of human behavior. This is because the organism acts as a whole, and its level of *general ability* (an ability to do many different things) is manifest in a variety of ways, some of which are more difficult to measure than others. Motor ability is one of these ways.

It is apparent that there is a powerful human quality closely tied in with general ability, and that is will. The function of will is to concentrate attention. A concentrated effort results from will power. It is this faculty that William James praised so highly, the faculty of voluntarily bringing back a wandering attention, over and over again. It is the prime characteristic that Maxwell Garnett, Spearman, McDougall, and many others have related directly to general ability.

What has all this to do with motor ability? From the time they are born certain children exhibit a remarkable aptitude for activities of a motor nature. As they grow older, they acquire physical skills with ease. They possess an aptitude for learning physical activities that can be called motor intelligence. They are "naturals" for athletic events, and very often, given good coaching, they become athletic champions. They appear to have the will and general ability to excel in motor activity. Specifically they have general motor ability, a qualitative component of physical capacity; demonstrated eye-hand coordination; power; strength; endurance; ability; speed; balance; eye-foot coordination; peripheral vision; and rhythm. In short, they are coordinated, as is the ballet dancer, the figure skater, or the master of broken field running.

The degree to which persons possess general motor ability and motor intelligence is the concern of physical education personnel. Just as human intelligence has an unknown potential limit, it is probable that motor intelligence also has an upper level difficult to estimate; with proper concentration, effort, and will power a given level of motor ability can be improved and made known through individual performance on standardized tests of motor ability.

This chapter will be devoted essentially to measures of general motor ability with a brief discussion of motor intelligence.

An Analysis of General Motor Ability

Ten elements of motor ability will be briefly discussed. These are the more common items associated with motor ability.

1. Strength and Endurance. Earlier when physical and motor fitness was being discussed it was shown that both quality and quantity of motor coordination is influenced by such elements as strength, muscular endurance, and cardiorespiratory endurance. Thus to be certain of an accurate change in motor ability, one should make certain that the physical fitness factor remains steady. With boys and girls, especially during the formative years, as strength increases, so does motor coordination. Willgoose studied adolescent boys aged twelve to eighteen years and found that speed for the dash and power for the standing broad jump is directly proportional to muscular strength.[1]

Strength index	Time for 50-yard dash (in sec)	Standing broad jump
Under 1000	7.9	
1001-1400	7.3	
1401-1800	7.0	6 ft 2 in.
1801-2200	6.8	6 ft 7 in.
2201-2600	6.6	7 ft 1 in.
2601-3000	6.3	7 ft 4 in.
over 3000	6.0	8 ft 2 in.

Also, at the college level it has been demonstrated that physical fitness and motor ability improve together.[2] Thus, both motor fitness and motor ability are essential to human development, but one is concerned primarily with organic qualities and the other with groups of skills.

2. Athletic Ability. In further analyzing motor ability one should note that, when motor behavior in team and individual sports is highly coordinated, the term *athletic ability* can be applied. The relationship, very often, between motor ability and athletic ability is a close one. Although it is theoretically possible for a person to have a high degree of general motor ability in handling the body and not possess many athletic skills, this is rare. School-age boys and girls tend to score high in athletic ability if they score

[1] Carl E. Willgoose, "Relationship of Muscular Strength to Motor Coordination in the Adolescent Period," *Journal of Educational Research*, 64:138–142, October, 1950.

[2] Carl W. Landiss, "Influences of Physical Education Activities on Motor Ability and Physical Fitness of Male Freshmen," *The Research Quarterly*, 26:295–307, October, 1955; see also Sidney Calvin, "Effects of Progressive Resistance Exercises on the Motor Coordination of Boys," *The Research Quarterly*, 30:287–298, December, 1959.

high in tests of motor ability. Top performers on motor ability tests will almost always be school athletes. They will also tend to be more mature for a given chronological age.[3]

3. Social Adjustment. There is abundant evidence to support a positive interrelationship between traits of personal and social adjustment and level of motor coordination.[4] This lends weight to the general ability thesis previously referred to.

When social adjustment, physical fitness, and motor ability approach the same plane, it is evidence in favor of the total body concept of behavior, a view expressed by Terman when he said that good things tend to go together.

4. Speed. Certainly speed and motor coordination are positively correlated. A short all-out sprint of 50–100 yards can be used as an index of speed. Because short bursts of speed are less affected by physical condition, they can be employed to rate a phase of general motor ability. This holds true up to the point where body fat and body density act as retarding elements; then tests of running and jumping are the most affected.[5]

Another test of speed, which also involves some agility, is the shuttle run. Here the subject runs back and forth several times between two points, sometimes starting and stopping as in the old-fashioned potato race.

The speed of performance is a rather stable motor element. Henry and Nelson, working with ten- and fifteen-year-old boys, found that individual differences in final skill are chiefly determined by initial skill.[6] About 50 per cent of the individual differences in motor learning were predictable from the prelearning performance. This seems to indicate that learning ability as such needs to be tested directly.

5. Reaction Time. One of the most involved elements of motor ability is personal reaction time. It is concerned with the ability to act (move) on a given stimulus. Two independent items are presented: simple reaction time and movement time. It has been found that a person may react quickly to a stimulus but move slowly, and vice versa. Because of this there has been a tendency to speak of these two items as unrelated.[7] Pierson, however, takes the view that there is a significant correlation between reaction time and

[3] Wilton Marion Krogman, "Maturation Age of 55 Boys in the Little League World Series, 1957," *The Research Quarterly*, 30:55–56, March, 1959.

[4] See especially the references at the end of this chapter by Jones, McCloy, and McCraw.

[5] R. P. Riendeau et al., "Relationship of Body Fat to Motor Fitness Test Scores," *The Research Quarterly*, 29:200–203, May, 1958.

[6] Franklin M. Henry and Gaylord A. Nelson, "Age Differences and Inter-relationship between Skill and Learning in Gross Motor Performance of Ten and Fifteen Year Old Boys," *The Research Quarterly*, 27:162–175, May, 1956.

[7] Franklin M. Henry, "Independence of Reaction and Movement Times and Equivalence of Sensory Motivation of Faster Response," *The Research Quarterly*, 23:43–53, March, 1952.

movement time.[8] Certainly the relationship of reaction time to athletic ability and the running speed of athletes has been well demonstrated.

6. **Balance.** The ability to maintain a specified position is a kind of static balance and may be compared with the ability to maintain a certain posture while moving, which is dynamic balance. Both kinds of balance have been related to general motor ability, with dynamic balance showing the strongest relationship.[9]

In a study of third- and fourth-grade girls it was found that balance and agility are significant elements in motor coordination.[10] This compared favorably with an earlier study of college women.[11] Specifically it was found that the capacity to balance objects and the speed of change of direction of arms and hands were strong factors. To the athletic coach this is not new knowledge. Balance and form have always been important items.

Estep carried out a study to investigate static equilibrium and success in gross motor activities.[12] Static equilibrium was measured by the Miles ataxiameter. This measure is a summation of anteroposterior and lateral postural sway. The subjects were selected on the bases of subjective ratings in motor ability in sport and rhythm, and A and C team skill classifications in the after-school sports program. This investigation supports the hypothesis that there is a positive relationship between balance and ability in gross motor activities; this was statistically significant at the 1 per cent and 5 per cent levels of confidence.

It is certainly a basic motor skill to maintain balance under varying physical circumstances. Gross and Thompson studied this factor and found significant relationships between dynamic balance and ability in swimming.[13] Slater-Hammel also found significant differences in balance scores. Using the Reynolds' Balance Test (static balance), he found the level of test performances to be, in descending order, varsity athletes, physical education majors, and liberal arts majors.[14]

[8] William R. Pierson, "The Relationship of Movement Time and Reaction Time from Childhood to Senility," *The Research Quarterly,* 30:227–235, May, 1959.

[9] Ruth I. Bass, "An Analysis of the Components of Tests of Semicircular Canal Function and of Static and Dynamic Balance," *The Research Quarterly,* 10:33–52, May, 1939.

[10] Frances Z. Cumbee, Margaret Meyer, and Gerald Peterson, "Factorial Analysis of Motor Co-ordination Variables for Third and Fourth Grade Girls," *The Research Quarterly,* 28:100–108, May, 1957.

[11] Frances Z. Cumbee, "A Factorial Analysis of Motor Co-ordination," *The Research Quarterly,* 25:412–428, December, 1954.

[12] Dorothy P. Estep, "Relationship of Static Equilibrium to Ability in Motor Activities," *The Research Quarterly,* 28:5–15, March, 1957.

[13] Elmer A. Gross and Hugh L. Thompson, "Relationship of Dynamic Balance to Speed and to Ability in Swimming," *The Research Quarterly,* 28:342–346, December, 1957.

[14] A. T. Slater-Hammel, "Performance of Selected Groups of Male College Students on the Reynolds' Balance Test," *The Research Quarterly,* 27:347–351, October, 1956.

7. Kinesthesis. Closely related to balance is the proprioceptive or kinesthetic consciousness of the body in movement. There is an awareness in space that some persons possess and others strive to attain. A good example of the latter is the tumbler or diver who loses himself during a flip and is never quite sure when to open up.

Efforts at measuring the kinesthetic function have been sporadic. Ragsdale, in discussing motor learning, points out that kinesthetic perception is fundamental and that it must be worked on to be developed.[15]

Motor accuracy is closely tied to kinesthesis, as are body position and direction of movement. With this in mind Wiebe carefully analyzed fifteen measures that seemed to be tests of kinesthesis and concluded that no single test is valid as a measure in this area.[16] Later, Scott arrived at approximately the same conclusion.[17] Much more research is needed before a test can be recommended for this element of motor ability.

8. Flexibility. There is a distinct relationship between motor ability and individual flexibility. But this must be considered on a specific basis; that is, flexibility for what purpose? More personal flexibility of structure may be needed for swimming than for wrestling. By and large, tension, the opposite characteristic of flexibility, is not helpful to motor performance. Yet there are exceptions. Each member of a small college soccer squad, for example, failed the flexibility part of the Kraus-Weber Tests and the team went on to win the national collegiate championship.

Tests of flexibility are covered separately at the end of this chapter.

9. Agility. This motor ability element is demonstrated by the capacity to change direction quickly and accurately. The Burpee test of World War II was designed to measure agility. From a standing position the subject drops to a squat with hands on the floor between the knees. The feet are then kicked backward to a front-leaning rest, then back to the squat position, and then to the standing position. The test is scored on the number of movements made in a twenty-second period.

Football, basketball, gymnastics and the high hurdles are all good examples of performances requiring agility. Mohr and Haverstick studied agility in sports and found significant associations between jumping ability and volleyball volleying and between agility and volleying.[18] Moreover, it seems that agility can be developed to a large degree, despite the fact that constitutional differences in body build are influential.

[15] C. E. Ragsdale, *How Children Learn the Motor Types of Activities,* 49th Yearbook of the National Society for the Study of Education. Chicago, University of Chicago Press, 1950.

[16] Vernon R. Wiebe, "A Study of Tests of Kinesthesis," *The Research Quarterly,* 25:222–228, May, 1954.

[17] M. Gladys Scott, "Measurement of Kinesthesis," *The Research Quarterly,* 26:325–341, October, 1955.

[18] Dorothy R. Mohr and Martha L. Haverstick, "Relationship between Height, Jumping Ability, and Agility to Volleyball Skill," *The Research Quarterly,* 27:74–78, March, 1956.

10. Psychological Capacities. In recent years it has become increasingly popular to study motor ability through psychological factors such as reaction time, emotional status, visual acuity, depth perception, attention, motivation, perception, and aptitude for learning in general.

Olsen, working with varsity athletes, intermediate athletes, and non-athletes, at Boston University, investigated the relationship between reaction time, depth perception, and visual span of apprehension as they relate to basketball, soccer, ice hockey, and baseball skills.[19] Although Olsen's results were not conclusive, the research warrants reading for the complications in this area are nicely brought out. Athletes were found to be superior to non-athletes in all tests and superior to intermediate athletes in reaction time. By and large, *human performance is a total body effort, and motor ability scores frequently reflect many physical and psychological capacities.* In all probability many more capacities are closely related to motor performance; only the intricacies of the variables cloud the true picture. As more research in these areas is carried on, the variables will be precisely delimited and will provide a clear picture of human behavior.

The Motor Intelligence Variable

What is native ability and why do some persons learn motor skills faster than others? The aptitude for learning motor skills is always present with great motor ability. One frequently wonders whether the observed state of general motor ability and over-all performance is the result of training in a number of physical activities or the result of an endowed motor intelligence. Most cited examples are probably a combination of both.

Motor intelligence is an aptitude for learning new motor skills. It is often referred to as *motor educability.*

Several factors complicate the study of motor learning. Maturation varies considerably within a chronological age group. Individual background skills of a person being studied influence learning aptitude scores. The amount of practice and instruction affect the acquisition and learning of motor skills. Moreover, physical traits, such as body build and visual perception, and other traits such as reaction time and span of apprehension are also significant.

Many of the motor ability measures significantly differentiate the fast and slow learners of physical skills. Smith found this to be true in studying the abilities of elementary school children.[20] It should be added however, that the validity of this study, and all studies concerned with the measurement of motor learning, is related in the final analysis to the method employed in measuring learning improvement. Brace made this clear in his research

[19] Einar A. Olsen, "Relationship between Psychological Capacities and Success in College Athletics," *The Research Quarterly*, 27:79–89, March, 1956.

[20] Jean A. Smith, "Relation of Certain Physical Traits and Abilities to Motor Learning in Elementary School Children," *The Research Quarterly*, 21:200–228, May, 1956.

on the rate of learning when he suggested that learning improvement should be divided by the possible gain rather than the initial score.[21] The possible gain is computed by subtracting the initial test score from the highest possible score that can be made on a test. McCraw studied eight test methods used to measure learning improvement and found that the most acceptable were those that related gain in motor ability to possible gain.[22]

Another interesting factor in the study of motor intelligence is that, when a person demonstrates a high level of motor ability, he is not only giving evidence of what he possesses at the moment; he is also showing *motor capacity.* Motor capacity is the limit of development of motor skill. A valid motor capacity score would predict how far a person could be expected to go in motor ability. It is, therefore, a factor of individual potentiality and may always be subject to speculation.

McCloy made an excellent beginning in this area and was followed by a number of others.[23] He acknowledged that motor capacity entails many things: environment, degree of constitutional endowment, and inherited characteristics. McCloy developed a test of motor capacity in which he considered four elements:

1. Size and maturity (using McCloy Classification Index to classify subjects)
2. Power measurement (vertical jump test)
3. Motor educability (intelligence factor measured by Iowa Brace Test—a stunt-type test)
4. Agility and coordination (Burpee test)

The scores obtained on these four tests were considered as measures of motor capacity. They were compared with the scores obtained on a battery of motor ability items (track and field events combined with chinning strength). By relating these two, motor capacity and motor ability, McCloy arrived at a General Motor Achievement Quotient. This GMAQ score became a kind of intelligence test for motor activity. In the Intelligence Quotient (IQ) chronological age expectations are compared with demonstrated mental age; in the GMAQ innate motor capacity is compared with developed motor ability.

Although it is relatively easy to set up a measure of motor ability that will show what a pupil can do at the moment, it is considerably more difficult to devise a test that will note a pupil's total capacity and indicate the motor limit over time. McCloy, as well as Brace, feels that the stunt-type items of motor educability, as expressed through such tests as the

[21] David K. Brace, "Studies in Motor Learning of Gross Bodily Motor Skills," *The Research Quarterly,* 17:242–253, December, 1946.

[22] L. W. McCraw, "Comparative Analysis of Methods of Scoring Tests of Motor Learning," *The Research Quarterly,* 26:440–453, December, 1955.

[23] Charles H. McCloy and Norma Young, *Tests and Measurements in Health and Physical Education,* New York, Appleton-Century-Crofts, Inc., 1954.

Johnson test and the Iowa Brace Test, are valid if they correlate well with improvement in age and poorly with strength, size, physiological maturity, and degree of individual skill. Immediately this presents a problem, for the interrelationship of these variables is so well established that it is almost impossible to hold certain variables under control while examining others. For example, the average boy and girl will show improvement on a physical test as they grow older, and very often the specific reason for this is improved skill and increase in size and strength. If it can be shown that the rate of learning improvement is greater than any influence attributed to these other factors, it may mean that the subject involved actually has greater motor intelligence or educability.

Continued research is needed to develop highly valid instruments to evaluate various aspects of motor performance as they relate to motor capacity and educability.

The tests that follow are divided into two groups: (1) motor ability and (2) motor intelligence. They are by no means complete in either variety or detail. They simply illustrate some of the research and useful measures available.

MEASURES OF MOTOR ABILITY

Standing Broad Jump

One of the basic large muscle motor activities employed over the years as a partial measure of general motor ability is the standing broad jump. There is considerable lever action in this event, which appears to be similar to body movements necessary in several other motor activities. This event, as well as the vertical jump, frequently correlates above .60 with motor ability scores. Several researchers, such as Brace, McCloy, Cureton, Rogers, and Scott, have advocated that jumping be included in test batteries designed to appraise motor ability.

Zimmerman studied the standing broad jump performance of skilled and nonskilled subjects and concluded that this event measures differences in performance of both classes of contestants.[24] Differences were found in angles of take-off and of landing and in extent and duration of specific joint actions. Moreover, there occurred a greater extent of movement in the skilled performances.

Sargent Vertical Jump

One of the motor ability tests that goes back to the early experiments of Dudley Allen Sargent is the jump and reach test, or vertical jump. It correlates well above .60 with the Strength Index test and the power events

[24] Helen M. Zimmerman, "Characteristic Likenesses and Differences between Skilled and Non-skilled Performance of Standing Broad Jump," *The Research Quarterly,* 27:352–362, October, 1956.

of track and field. Where practice is permitted first, reliability coefficients have been reported above .85.[25] It can be used with elementary and junior high school boys and girls. In fact, they usually enjoy taking the test and comparing themselves as they do in other stuntlike events.

In this event the subject demonstrates the ability to develop power in relation to his weight. This is accomplished by having the subject stand on a hard, level surface with one side to a vertical wall. With both feet flat on the floor he reaches as high as he can with the hand nearest the wall. This

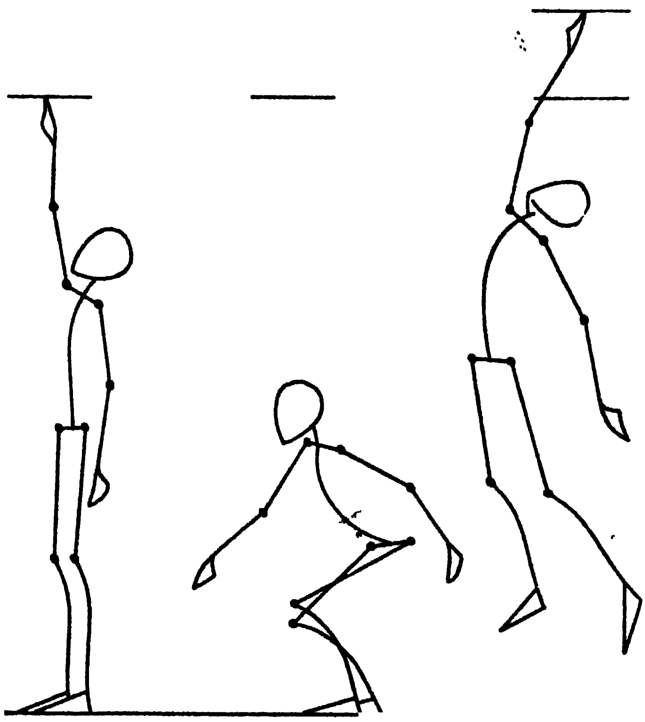


Fig. 11-1. The Vertical Jump Test.

height can be marked by the subject with a piece of chalk. The subject then moves 2 inches away from the wall, crouches until his upper and lower legs form about a right angle, and jumps, again reaching up the wall as high as possible (see Figure 11-1). No extra hop or step is taken before the jump. The second height is marked with the chalk, and the difference between the maximum standing reach and the maximum jumping reach is recorded to the nearest inch. The best jump of three trials is recorded as the score. A 1-foot wide, 5-foot long sheet of brown paper glued to a flat wall

[25] Charles H. McCloy, "Recent Studies in the Sargent Jump," *The Research Quarterly*, 3:235, May, 1932; see also Betty A. Pachew, "Effectiveness of Warm-up Exercises on Junior High School Girls," *The Research Quarterly*, 30:202–213, May, 1959.

may also be used. The bottom edge of the paper should be about 6 feet above the floor, and it should be ruled with horizontal lines about 1 centimeter apart. By placing a chair to one side of the paper the instructor may readily note the height of the jump.

In order to reduce error of measurement McCloy developed a leapmeter. This is a device where a skull cap, from which an attached string goes vertically to a lever arm and wall pulley, is placed on the subject's head. When the subject is standing, a mark is made at the end of the weighted string as it hangs on the wall. At the height of the jump the end of the string will drop, and the difference in the two reachings may be calculated by noting the wall measurements.

In the Larson Muscular Strength Test the vertical jump is combined with pull-up and push-up scores. This test is a practical one for predicting general motor ability. For a discussion of this test, see Chapter 7.

Rogers' Strength Index

The strength index (SI) as a measure of muscular strength and power correlates well with general motor ability. Rogers' original work on the physical fitness index was done in the area of general athletic ability, where SI scores were related specifically to a combination of scores on the bar vault, running broad jump, running high jump, and 100-yard dash.[26] Here a correlation of .76 was obtained. This was improved to .81 when the SI was related to the physical skills of throwing baseballs and footballs for distance, basketball foul throws, 8-pound shot-put, running high jump, standing broad jump, and time in a two-lap running event. Other researchers have found that McCloy's revision of the strength index is a very good index of ability to learn new or unfamiliar motor activities.[27]

Numerous studies on the equalization of teams for intramural athletics demonstrate beyond a doubt the close relationship between general motor (athletic) ability and strength status. This is just as true for girls and women as it is for boys and men. Anderson, for example, showed that the athletic performance of high school girls correlated .65 with McCloy's strength index.[28] Carpenter proved, among other things, that both strength and speed are significant elements in the athletic performance of women.[29] This finding is in keeping with current growth and development research, especially when motor coordination is appraised during the adolescent period.

[26] Frederick Rand Rogers, *Physical Capacity Tests in the Administration of Physical Education*; New York, Bureau of Publications, Teachers College, Columbia University, 1925.

[27] Elmer A. Gross, Donald C. Greisel, and Alan Stull, "Relationship between Two Motor Educability Tests, a Strength Test, and Wrestling Ability after Eight Weeks Instruction," *The Research Quarterly*, 27:395–401, December, 1956.

[28] Theresa Anderson, "Studies in Strength Testing for High School Girls," *The Research Quarterly*, 8:69–74, October, 1937.

[29] Aileen Carpenter, "A Critical Study of Factors Determining Strength Tests for Women," *Research Quarterly*, 9:3–6, December, 1938.

This point will be referred to in more detail in Chapter 14, which treats the topic of classification of pupils for physical education activities.

Oberlin College Test

One of the oldest collegiate tests for appraising general motor ability of men students is a ten-item test developed at Oberlin College, in Ohio.[30] A standard performance for each event is set up. Meeting the standard for each event scores 10 points toward a perfect score of 100 points. The ten tests are as follows:

Test item	Type of performance	Standard
1. Running	Two laps of indoor track (176 yd)	24 sec
2. Jumping	Running high jump	4 ft 10 in.
3. Vaulting	Over low horizontal bar from standing position	47 in. high
4. Climbing	20-ft rope from kneeling start	12 sec
5. Lifting and pulling	Two backward circles on high horizontal bar, from hanging start with arms extended	Continuous movement for each
6. Pushing	Parallel bar dips	Ten times
7. Throwing	Throw for accuracy; baseball at 18-in. circle at 60 ft	Three hits out of five throws
8. Swimming	100-yd free style	1 min 45 sec
9. Tumbling	Handspring on mat	Satisfactory
10. Balancing	Handstand within 4-ft circle	10 sec

McCloy's General Motor Ability Tests

To McCloy motor ability is the result of innate characteristics plus motor learning. It is a developed capacity for motor activities. By experimentally combining muscular strength with track and field events a formula was obtained for calculating the *General Motor Ability Score* (GMAS)[31]:

GMAS = .1022 (total track and field points) + .3928 (chinning strength)[32]

For boys the track and field events included (1) a sprint in seconds varying from 50 to 199 yards; (2) a standing or running broad jump; (3) a running high jump; and (4) a weight throwing event, shot-put, or basketball or baseball throw for distance. Scoring is done on McCloy's scoring tables.

A similar test for girls consists of the sprint, broad jump, and throw for

[30] David K. Brace, *Measuring Motor Ability*, New York, A. S. Barnes and Company, 1927.

[31] Charles H. McCloy and Norma D. Young, *op. cit.*, p. 207

[32] According to McCloy's strength index formula:

Dipping or chinning strength = 1.77 (weight of subject) + 342 (chins or dips) − 46

distance. McCloy's scoring tables are used and applied to the following formula for girls.

$$GMAS = .42 \text{ (track and field points)} + 9.6 \text{ (number of chins)}$$

McCloy also developed universal scoring tables for use with twenty-seven track and field events together with floor push-ups and chins. Scores can be compared with each other. By combining the universal scoring tables and his classification index, McCloy created the *Athletic Quotient*, the percentage that the particular motor performance bears to the average performance of secondary school boys of the same age, height, and weight. It can be used for classifying teams or for creating homogeneous categories of pupils.

As previously indicated in a discussion of motor intelligence, McCloy worked out a test of *motor capacity* (General Motor Capacity Score GMCS), a measure of potential motor ability for developing motor skills. This is an inborn capacity. The test was validated against a battery of track and field tests plus chinning strength. The Iowa Brace Test of motor educability, the Burpee test, the vertical jump score, and McCloy's classification index are combined into this test of motor capacity. A *Motor Quotient* (MQ) can be obtained that is similar to the IQ score in meaning. A raw score on a motor capacity test (GMCS) is divided by a norm based on the classification index (boys or girls for a given age). The tables, which are found in the cited reference, consider a per cent relationship of actual ability to the standard ability. An MQ score of 70, therefore, would show that a subject's performance is only 70 per cent of his potential capacity. Likewise an MQ of 100 would be average for a boy or girl of a certain age, height, and weight. The boy scoring 125 would have little trouble learning new motor skills, but a pupil scoring 78 would be expected to learn new motor skills very slowly.

When motor capacity (GMCS) is related to motor ability (GMAS) a *General Motor Achievement Quotient* (GMAQ) results.

$$\frac{GMAS}{GMCS} \times 100 = GMAQ$$

This is achievement. It represents how well a person does in terms of his potential capacity. It will rise or fall with current ability in motor skills.

Cozens Test of General Athletic Ability

Cozens developed his test as a means of finding strengths and weaknesses in physical achievement in instructional classes of physical education.[33] An extensive study of forty motor activities was made with the help

[33] F. W. Cozens, *The Measurement of General Athletic Ability in College Men*, Eugene, Ore., University of Oregon Press, 1929; also *Achievement Scales in Physical Education Activities for College Men*, Philadelphia, Lea & Febiger, 1936

of physical education judges. From this seven tests were selected to measure the elements of general motor ability. Each test was weighted by a multiplier according to the degree to which it alone contributed to general motor ability. The seven tests, together with the specific element measured and its weighting factor, are as follows:

Test item	Element measured	Multiplier
1. Dips on parallel bars	Shoulder-girdle and arm strength	0.8
2. Baseball throw for distance	Shoulder-girdle and arm coördination	1.5
3. Football punt for distance	Foot-eye, arm-eye, hand-eye coordination	1.0
4. Standing broad jump	Leg strength, power, and flexibility	0.9
5. Bar-snap for distance	Body ability, control, and coordination	0.5
6. Dodging run	Leg speed	1.0
7. Quarter-mile run	Endurance	1.3

The validity and reliability of these tests has run high over the years, making the Cozens battery one of the most useful. Another reason they have been so well received is that they can be applied strictly on an individual basis, which includes differences in physique. Norm charts are set up to take into account nine height-weight divisions (see Achievement Scales in reference cited):

Tall slender	Medium slender	Short slender
Tall medium	Medium medium	Short medium
Tall heavy	Medium heavy	Short heavy

Short men are those 5 feet 6¾ inches or less; tall men are those 5 feet 10½ inches and over. A man is scored according to his height-weight division for each event. Raw scores are changed to sigma scale scores from the achievement tables and multiplied by the multiplier. Total scores are related to a classification item where the performance is rated as superior, above average, average, below average, and inferior. From this classification special physical education assignment can be made so the subject can begin working to show improvement at a later date.

The test items are administered as follows:

1. **Dips.** The subject starts in a support position on the ends of the parallel bars. He lowers the body until the upper arm is parallel with the floor or beyond. He then pushes up until the arms are straight. This is repeated as many times as possible.

2. **Baseball Throw for Distance.** After a one-minute warm-up the subject throws the ball as far down the field as he can. He is allowed a 15-foot run, and the best of three throws is scored. A 220-yard field is divided into ten sections to facilitate distance measurements for this event.

3. **Football Punt for Distance.** The best of three punts is recorded. A 15-foot restraining area is allowed, leading to a field 165 yards long and 125 yards wide, with 5-yard line markings.

4. **Standing Broad Jump.** Standard method.

5. **Bar-snap for Distance.** The subject grasps the horizontal bar (set 4 feet 6 inches above the floor), swings underneath with the feet close to the bar, shoots the feet out and upward. archs the back, and lets go the bar at the proper instant to throw the body to the feet a distance beyond the bar. Three trials are permitted with the greatest distance being recorded as the score. Practice is allowed for this event.

6. **Dodging Run.** A special layout is needed for this event; see Figure 11-2. The subject begins at the starting line (A) and runs around the

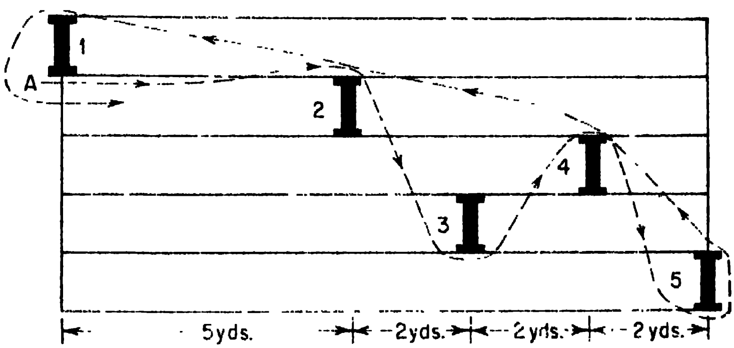


Fig. 11-2. Cozens Dodging Run.

hurdles over the prescribed course for two complete trips. Although the subject may jog over the course to familiarize himself, only one trial is allowed in the test. Time is recorded in seconds from the word "go."

7. **Quarter-mile Run.** One trial is permitted; it is scored to the nearest second.

Larson's Motor Ability Test

In the interest of homogeneous classification of motor ability Larson developed an indoor test and an outdoor test.[34] They were derived from twenty-five motor ability elements, are easy to administer, and have satisfactory validity and reliability.

The Indoor Test

1. *Dodging Run:* Performed over the same course and exactly as in the Cozens Test of General Athletic Ability (see Figure 11-2).
2. *Chinning:* From high horizontal bar.
3. *Dipping:* From parallel bars as in Cozens test.

[34] Leonard A. Larson, "A Factor Analysis of Motor Ability Variables and Tests for College Men," *The Research Quarterly,* 21:499–517, October, 1941.

4. *Vertical Jump:* Standard procedure using calibrated wall board.
5. *Bar Snap:* Same procedure as in Cozens test.

The Outdoor Test

1. *Baseball Throw for Distance:* 12-inch inseam softball used.
2. *Bar Snap:* Same as Indoor Test.
3. *Chinning:* Same as Indoor Test.
4. *Vertical Jump:* Same as Indoor Test.

In scoring the test the raw scores for each event are converted to T scores. These are converted to weighted T scores with a multiplier. The weighted scores for the indoor or outdoor test are then totaled to provide an index score. Norms for the outdoor test appear in Table A-15, Appendix A. Total scores for the outdoor test can be used to classify men as follows:

Excellent motor ability	= 728–up
Good motor ability	= 642–727
Average motor ability	= 556–641
Poor motor ability	= 470–555
Very poor motor ability	= 469–down

Carpenter Motor Ability Test

Desiring to promote motor ability in primary school children, Carpenter experimented widely and built a test for boys and girls consisting of three parts:[35]

1. *Broad Jump:* Standard Procedure. Recorded in inches.
2. *Shot-put:* A 4-pound shot, put according to standard procedure and recorded to the nearest foot.
3. *Weight of Subject:* Recorded to nearest pound.

This simple test is based in part on research carried on by Stansbury a year earlier, in which a test was built for high school boys by relating the broad jump, body weight, and the results of the 8-pound shot-put to motor performance.[36]

In scoring, two equations are employed to predict general motor ability.

Boys: General motor ability score = (broad jump) + 2.3 (shot) + (weight)

Girls: General motor ability score = .5 (broad jump) + 3 (shot) + (weight)

[35] Aileen Carpenter, "The Measurement of General Motor Capacity and General Motor Ability in the First Three Grades," *The Research Quarterly,* 13:444–465, December, 1942; see also "Strength Testing in First Three Grades," *The Research Quarterly,* 13:332, October, 1942.

[36] Edgar Stansbury, "A Simplified Method of Classifying Junior and Senior High School Boys into Homogeneous Groups for Physical Education Activities," *The Research Quarterly,* 12:765–776, December, 1941.

Norm tables for the Carpenter test are available in Tables A-4 and A-5, Appendix A. They are used in conjunction with McCloy's Classification Index for high school boys [20 (age) + 6 (height) + weight]. This is done despite the fact that height does not appear to be a factor in the lower grades. When the classification index is figured, it is observed in the scoring table to see what norm is called for. A norm of 100 represents satisfactory motor ability for a certain sex, age, and weight.

To calculate the *Physical Efficiency Index* the motor ability score achieved by the subject is multiplied by 100 and divided by the norm from the norm table. A norm of 100 represents satisfactory motor ability for a certain sex, age, height, and weight.

Example:

Girl: Age = 7 years
Height = 48 inches
Weight = 52 pounds
Broad jump = 45 inches
Shot-put = 8 feet

General motor ability score = .5(45) + 3(8) + 52 = 98.5

Class. index = 20 (age) + 6 (height) + weight = 140 + 288 + 52 = 480

Norm (for 480) = 94.44

$$\frac{\text{General motor ability score } 98.5 \times 100 = 9850}{\text{Norm } 94.44} = 104$$

Here the PEI of 104 indicates that this elementary school girl has just above what is expected of her in motor ability, for she is just above the norm of 100.

This test warrants greater use, for it is easy to administer and interpret. Carpenter has worked out regression equations whereby McCloy's Elementary School Age Classification Index formula can be substituted. This involves another step and makes the scoring process a little more time-consuming.

Humiston Motor Ability Test

Another test, used specifically with college women, is one developed before World War II by Humiston to measure present status.[37] The test, based on seven items, has a coefficient of validity of .92 when the items are administered separately. Its use is to classify pupils for physical education and intramural play.

A rather complicated obstacle course is set up in the gymnasium or similar area where the girl has to execute a dodging run, a sideward roll on a mat, climb over a box, turn in a circle and proceed between a number

[37] Dorothy Humiston, "A Measure of Motor Ability in Women," *The Research Quarterly*, 8:181–185, May, 1937.

of barriers, climb a ladder, throw a basketball over a rope, and sprint a short distance.

An adaptation of the Humiston test has been made by substituting ten motor items in the obstacle course for the original seven and calling it the Olympic Motor Ability Test.[38] The adaptation appears valid for high school girls and perhaps a little easier to execute than the original test. Scoring is accomplished by recording the number of seconds it takes to complete the course of obstacles. Percentile norms are available in the Humiston reference, but not for the Olympic test.

Barrow Motor Ability Test

The purpose of the Barrow test is to appraise easily general motor ability of college men for purposes of classification, guidance, and measurement of achievement.[39] Barrow employed expert opinion in the validation process. Twenty-nine items were used to measure eight factors of motor ability. The selected tests were administered to 222 college men. Through use of the Doolittle technique multiple correlations and regression equations were computed for several combinations of items. Two test batteries are recommended for use. Norm tables and profile cards are available in a test manual distributed by the author. Scoring tables appear in the cited references.

Test items for the Barrow test consist of (1) standing broad jump; (2) softball throw; (3) zigzag run; (4) wall pass; (5) medicine ball put; and (6) 60-yard dash.

A short form of the test is the three-item battery, which consists of the standing broad jump, medicine ball put, and zigzag run. It has high validity, reliability, and objectivity and appears to be a quick method of pupil classification. It has the further advantage of being completely suitable for indoor appraisals.

Newton Motor Ability Test

In this high school test for girls Powell and Howe tried out several different motor test items in a number of ways before arriving at a three-item battery consisting of the standing broad jump, a hurdle race, and a "scramble" test.[40] Several acceptable criteria were used to validate the tests, including an analysis of eighteen tests of power, strength, speed, and coordination and rating by a jury of competent judges.

In the initial study the test results appeared to be about 3½ times as

[38] Shirley J. Kammeyer, "Reliability and Validity of Motor Ability Test for High School Girls," *The Research Quarterly*, 27:310–315, October, 1956.

[39] Harold M. Barrow, "Test of Motor Ability for College Men," *The Research Quarterly*, 25:253–260, October, 1954; test manual available from author at Wake Forest College, Wake Forest, N.C.

[40] Elizabeth Powell and E. C. Howe, "Motor Ability Test for High School Girls," *The Research Quarterly*, 10:81–88, December, 1939.

good as the Rogers' PFI in predicting motor ability. This should be expected to some extent, as the PFI is more of a fitness and endurance measure and the SI is a better measure of motor ability. In women, however, strength does not appear to be as significant as it is with men in predicting all-round motor performance. Anderson noted this when she found that the motor ability of high school girls was related more to the vertical jump than to the strength index.[41]

Achievement scales are available in the cited reference for each of the test items. Also, total points and decile scores for the tests are on hand to help in more effective judgment of motor performance. The tests are as follows:

1. **Broad Jump.** Three trials are permitted with the best trial recorded to the nearest inch.

2. **Hurdles.** One trial is given to run over six hurdles set at prescribed distances apart (run six hurdles in one direction, turn around an Indian club, and six hurdles back to starting line for finish) measured to the nearest fifth of a second.

3. **Scramble.** The subject lies flat on her back with feet against the wall. Arms are raised sideward to shoulder level with palms down on the floor. On a "go" signal the subject gets to her feet quickly and runs to a bell 10 feet away, taps it twice, and returns to the floor position, where she claps her hands on the floor twice. This whole process is repeated four times and measured to the nearest fifth of a second from the word "go."

Scott Motor Ability Test

Although originally validated on college women, this test can be used with high school girls as a measure of general ability.[42] Scott carried on an extensive analysis of the motor behavior of women, investigating all sports skills and combination of skills. From this research she proposed a five-item battery as follows:

1. **Basketball Throw for Distance.** The purpose is to measure shoulder girdle strength and coordination. The ball is thrown as far as possible from behind a restraining line. The best of three trials is recorded to the nearest foot.

2. **Four-second Dash.** In order to test the speed of movement the subject, on the starting signal, runs along a straight lane 85 to 90 feet long and 4 feet wide. The lane is carefully marked with a white line or a series of white cards placed a yard apart. After four seconds has elapsed, the starter blows a whistle and the assistant quickly takes note of where the runner is in the running lane. This distance from the beginning is scored in yards.

[41] Theresa W. Anderson, "Weighted Strength Tests for the Prediction of Athletic Ability in High School Girls," *The Research Quarterly*, 7:27–40, March, 1936.

[42] M. Gladys Scott, "The Assessment of Motor Abilities of College Women through Objective Tests," *The Research Quarterly*, 10:63–83, October, 1939; see also "Motor Ability Tests for College Women," *The Research Quarterly*, 14:402–405, December, 1943.

Only one trial is permitted, and the girl may use any starting position she choses.

3. **Wall Pass.** To test skill in handling and controlling a ball the subject stands behind a restraining line 9 feet away from a flat wall (at least 8 feet square) and rebounds the ball off the wall as quickly as possible. The object is to score as many hits as possible in a fifteen-second period. A dropped ball is recovered, and rebounding continues. Although some practice throws are permitted, only one trial is recorded.

4. **Standing Broad Jump.** Leg power is appraised by the usual procedure; the best of three trials is recorded to the nearest inch.

5. **Obstacle Race.** This test is designed to measure a combination of speed, ability, and general body coordination. The subject lies on her back with heels at the starting line. On the "go" signal she gets up quickly, and runs at

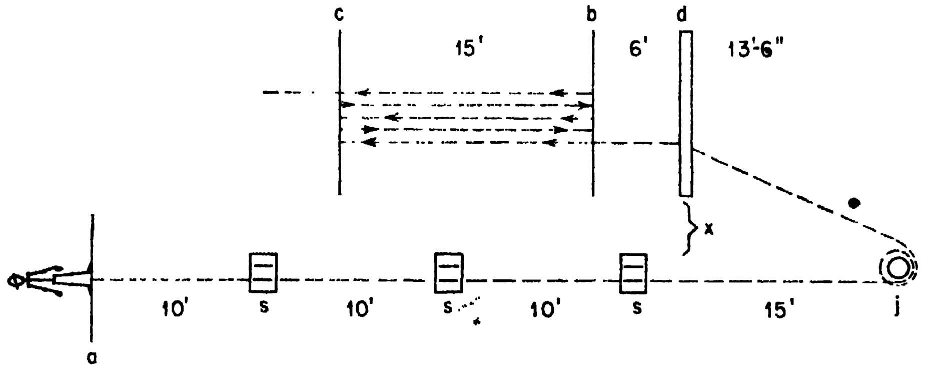


Fig. 11-3. Markings for Scott's Obstacle Race. *a*, Starting line; *b*, line for shuttle; *c*, finish line; *d*, boom, 18 inches high; *j*, jump standard; *s*, spot on floor, 12 by 18 inches; *x*, distance from end of boom to line of inner side of spots, 4 feet 4 inches.

top speed toward the first obstacle (jump standard, see Figure 11-3). As she approaches each rectangle on the floor, she must step on it with both feet. Upon reaching the jump standard, the subject runs around it twice and heads toward the crossbar (boom) to go under it. Completing this, she runs to line *c*, turns about, and runs back to line *b*. This is done three times; when the subject crosses line *c* for the third time, she has finished the obstacle race. Time is recorded to the nearest tenth of a second, from the "go" signal until line *c* is crossed for the last time.

In the interest of time the wall pass and the four-second dash can be omitted from the test battery without seriously affecting the validity of the measure.

In calculating the general motor ability score the raw scores for each test item are converted to T scores. The T scores are then summed and divided by the number of tests given to obtain an average. A high T-score average indicates above-average general motor ability. See Table A-24, Appendix A.

Example

High School Girl: Three-item Test

	Raw score	*T score*
Obstacle race..........	19.0 seconds	70
Broad jump............	86 inches	68
Basketball throw.. .	63 feet	73
Average T score..................	(3) 211	= 70+

MEASURES OF MOTOR INTELLIGENCE

Iowa Brace Test

In his early studies of motor educability in the mid-twenties, Brace experimented with twenty self-testing stunts, which were scored on a pass or fail basis.[43] At that time no consideration was given to gross differences in body structure and age.

Subsequently, McCloy, of Iowa, experimented with over forty stunts and gradually reduced them to twenty-one.[44] Of this number ten were in the original Brace battery. The tests selected by McCloy, however, were considered better measures of native motor ability because they were relatively free from strength elements and learned movements. Obviously a test of motor intelligence or educability should consist of motor items not readily improved upon by practice. As one might suspect, this is where the validity of motor intelligence tests is challenged, for it seems possible to improve performance in almost any motor area when specific practice is engaged in.

From the Iowa improvement of twenty-one stunts a total of six test batteries of ten stunts each was put together for grades four to six, junior high school, and senior high school. On each level there is a set of stunts for boys and a set for girls.

In the stunts that follow absolutely no practice is allowed. Two trials are given on each test; they are scored on pass or fail. Success on the first trial scores two points, on the second trial one point. No points are scored for a failure. With 20 points representing the highest possible score, it is not difficult to compare subjects. McCloy prepared T-score tables, which are found in the cited reference.

Because the validity of these tests as measures of motor educability is questionable, only a brief description of the stunt items will be included here:

Test 1. One Foot—Touch Head.[45] Standing on the left foot, bend forward and place both hands on the floor. Then raise the right leg and stretch

[43] David K. Brace, *Measuring Motor Ability.*

[44] Charles H. McCloy, "An Analytical Study of the Stunt Type Tests as a Measure of Motor Educability," *The Research Quarterly,* 8:250–265, October, 1937; see also Charles H. McCloy and Norma D. Young, *op. cit.,* part III.

[45] Originally included in the Brace test: Tests, 1, 3, 5–8, 14, 16, and 21.

it back. Touch the head to the floor, and regain the standing position without losing balance.

Test 2. Side-leaning Rest. From a sitting position on the floor with the legs straight out and the feet together, place the right hand on the floor behind the body. Then turn to the right and assume a side-leaning rest position, resting on the right hand and right foot. Raise the left arm and hold this position for five counts.

Test 3. Grapevine. From a standing position with both heels tight together, bend down and extend both arms down between the knees, around behind the ankles, holding the fingers together in front of the ankles without losing balance.

Test 4. One-knee Balance. Kneel down on one knee, with the other leg raised from the floor and arms stretched out to the side. Hold the position of balance for five counts.

Test 5. Stork Stand. While standing on the left foot, hold the bottom of the right foot against the inside of the left knee. With the hands on the hips and both eyes closed hold the position for ten seconds without moving the left foot about on the floor.

Test 6. Double Heel Click. Jump into the air, clap the feet together twice, and land with the feet any distance apart.

Test 7. Cross-leg Squat. With arms folded across the chest, cross the feet and sit down cross-legged. The object is to arise without unfolding the arms or having to move the feet to regain balance.

Test 8. Full Left Turn. From a standing position with feet together, jump into the air, make a full turn to the left, and land on the same spot. This is done without losing balance or moving the feet after landing on the floor.

Test 9. One Knee—Head to Floor. While kneeling on one knee, with the other leg stretched out behind and not touching the floor, extend the arms to the side parallel to the floor. Then bend forward and touch the head to the floor. Raise it without losing balance.

Test 10. Hop Backward. Standing on either foot, close the eyes and take five hops backward without dropping the other foot or opening the eyes.

Test 11. Forward Hand Kick. Jump in the air, lifting both legs forward, and touch the toes with both hands before landing. Knees cannot be bent more than 45 degrees.

Test 12. Full Squat—Arm Circles. From a full squat position with the arms out to the sides, wave the arms so the hands make a circle about 1 foot across, "jiggling" up and down for ten counts.

Test 13. Half-turn Jump—Left Foot. While standing on the left foot, jump a half turn to the left, maintaining balance.

Test 14. Three Dips. With the hands placed on the floor, arms straight, extend the feet back along the floor until the body is straight. Perform three push-ups with only the chest touching the floor each time.

Test 15. Side Kick. Jump upward from the right foot, throwing the left foot sideways to the left. Strike the feet together in the air, and land with the feet apart.

Test 16. Kneel, Jump to Feet. From a kneeling position on both knees, with toes extended out flat behind, swing the arms and jump to the feet without rocking back on the toes or losing balance.

Test 17. Russian Dance. From a low squat position with one leg stretched forward, hop and reverse legs as in the Russian dance step. This is performed twice without losing balance.

Test 18. Full Right Turn. Starting with both feet together, swing the arms and jump into the air, executing a full turn to the right. The subject must land on the same spot without losing balance or causing the feet to move after they strike the floor.

Test 19. The Top. From a sitting position on the floor, the arms are placed between the legs and under and behind the knees to grasp the ankles. A roll-around is executed rapidly to the right with the weight first over the right knee, then the right shoulder, then over the back, then left shoulder, then left knee. The subject then sits up facing in the opposite direction from which he started. This is repeated with the subject ending up in the same direction from which he originally started.

Test 20. Single Squat Balance. From a squat position on either foot the other leg is stretched forward off the floor. With hands on hips, the position is held for five counts without losing balance.

Test 21. Jump Foot. Holding the toes of either foot in the opposite hand, jump up with the free foot and over the foot that is being held without letting go or losing balance.

The breakdown of the ten tests according to age level is as follows:

Age level	Test number																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Boys Grades 4–6		x	x	x			x	x		x	x		x			x	x				
Junior high school	x	x	x			x						x	x	x		x	x		x		
Senior high school	x		x		x						x			x	x	x	x			x	x
Girls Grades 4–6	x		x			x		x		x	x				x	x		x	x		
Junior high school	x	x									x	x	x		x	x	x		x	x	
Senior high school		x	x				x		x		x					x	x	x	x	x	

As might be suspected, all these tests are affected considerably by practice. To overcome this difficulty, McCloy suggests that the stunts be given no oftener than once every three years for the same group of boys or girls.

In correlating ability to learn various sports skills with scores on the Brace and Iowa Brace tests, it was found that a low correlation of agreement exists but certainly not high enough for prediction purposes. In fact, in only one area did Brace obtain a correlation of .50, and that was the ability to learn rhythmic skills.[46] Moreover, later on Brace concluded that his test did not measure motor learning to a satisfactory degree and neither did the Iowa Brace test.[47]

In short, the test is a good one to study when exploring the area of motor intelligence as it relates to motor learning, but it is limited in its use.

Johnson Test of Motor Educability

A few years after the Brace test appeared, Johnson designed a motor educability measure that could be used to group pupils with a similar degree of ability for learning motor skills.[48]

In this test a series of ten stunt-type exercises are performed on an especially marked 5- by 10-foot gymnasium floor mat. Johnson's validity coefficient was low (.69), but Koob compared the test scores of junior high school boys with their performance in learning ten different tumbling stunts and obtained a correlation of agreement of .95.[49] Although stunt-type tests are frequently lacking in high objectivity and reliability, Johnson reported an r of .97. Gire and Espenschade, working with high school girls, only obtained a .61 coefficient of correlation.[50] Other research is in agreement with these findings. Hatlestad, for instance, reviewed the test in her study of motor educability in college women and concluded that greater objectivity is needed before it could have widespread use.[51]

The ten items that appear in the Johnson test are (1) straddle jump; (2) stagger skip; (3) stagger jump; (4) forward skip (holding opposite foot from behind); (5) front roll; (6) jumping half turns, right or left; (7) back roll; (8) jumping half turns, right and left alternately; (9) front and back roll combination; (10) jumping full turns.

[46] David K. Brace, "Studies in the Rate of Learning Gross Bodily Motor Skills," *The Research Quarterly*, 12:181–186, May, 1941.

[47] David K. Brace, "Studies in Motor Learning of Gross Bodily Motor Skills."

[48] Granville B. Johnson, "Physical Skill Tests for Sectioning Classes into Homogeneous Units," *The Research Quarterly*, 3:128–134, March, 1932.

[49] Clarence G. Koob, "A Study of the Johnson Skills Test as a Measure of Motor Educability," unpublished master's thesis, State University of Iowa, Iowa City, Iowa, 1937.

[50] Eugenia Gire and Anna Espenschade, "The Relationship between Measures of Motor Educability and the Learning of Specific Skills," *The Research Quarterly*, 13:43–52, March, 1942.

[51] L. Lucile Hatlestad, "Motor Educability Tests for Women College Students," *The Research Quarterly*, 13:10–17, March, 1942.

Johnson-Metheny Test

In analyzing the Johnson test with an eye toward some improvements, Metheny discovered that a limited number of the stunt items in combination correlated quite well with learning ability in boys and girls.[52] This prompted further investigation, the result of which indicated a correlation of .98 between Tests 5, 7, 8, and 10 and the total Johnson test score for boys. Moreover, the same four tests correlated quite well (.93) with learning tumbling stunts. In measuring girls, Tests 5, 7, and 8 appeared suitable for substitution for the total Johnson test, as they correlated .87 with it.

By employing only three items for girls and four for boys from the original tests, Metheny was able to create a relatively simple target (Figure 11-4). The markings are made on a mat canvas 15 feet long. A lane 24 inches wide is marked down the center of the mat; it is further divided into two equal narrow lanes by a center line and into ten equal parts lengthwise by parallel lines drawn every 18 inches. These lines are alternately 3 inches wide and ¾ inch wide, with the centers 18 inches apart.

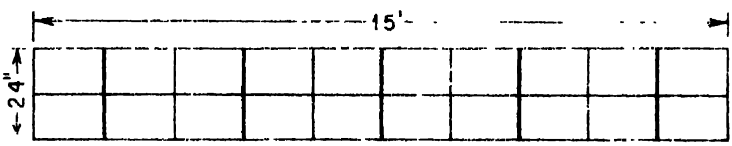


Fig. 11-4. Johnson-Metheny Floor Canvas Markings.

When this canvas target is placed over gymnasium mats, with its ends tucked under the mats and pulled taut, it is ready to be used. Tests 5, 7, 8, and 10, which make up the Johnson-Metheny test, are performed as follows:

A. (Test 5) **Front Roll.** From a standing position off the chart at the end of the 24-inch lane, the subject performs two front rolls. The first is done within the limits of the first half of the lane, not going beyond the middle 3-inch line. The second roll is done within the limits of the second half of the lane, never touching or overlapping the lanes. Five points are given for each roll; two points are deducted for overreaching side line left or right for each roll. One point is deducted for overreaching the end limits on each roll; and five points for inability to perform a true roll.

B. (Test 7) Back Roll. The subject performs two back rolls the entire length of the 24-inch lane. One roll is performed in each half as in the previous test. Scoring is also the same.

C. (Test 8) Jumping Half Turns, Right and Left Alternately. With both feet on the first 3-inch line, the subject jumps with both feet to the second 3-inch line, executing a half turn in the opposite direction. This is

[52] Eleanor Metheny, "Studies of the Johnson Test as a Test of Motor Educability," *The Research Quarterly*, 9:105–114, December, 1938.

continued the length of the mat, with alternating directions of rotation. Ten points is a perfect score. Two points are deducted for each jump where the subject does not land with both feet on the 3-inch line, or turns the wrong way, or both.

D. (Test 10) **Jumping Full Turns.** Beginning from a position at the end of the target with the feet outside the lane, the subject jumps with the feet together to the second rectangular space, executing a full turn with the body right or left. This is continued down the mat with each jump; that is, full turns are made, rotating in the same direction and landing on both feet in every second rectangular space. Ten points is a perfect score. Two points are deducted if the subject fails to land on both feet, oversteps the rectangle, turns too far or not far enough, or loses balance prior to starting the next jump.

A perfect score for the total test is 40 points.

Adams Sport-type Test of Motor Educability

In studying motor educability Adams suggested that the ability to learn stunts might be different from the ability to learn sport skills, and that they may not be highly correlated.[53] Working with college men, Adams selected four sport-type tests from a group of forty-nine tests. They are as follows:

1. **Wall Volley Test.** Standing three feet away from a flat wall, the subject volleys a volleyball so it strikes the wall above a line drawn 10½ feet above the floor. Each trial is scored on the basis of the number of consecutive volleys up to ten. The sum of the scores on seven trials represents the total score.

2. **Lying Tennis Ball Catch.** Here the subject, lying flat on his back holding a tennis ball, tosses the ball 6 feet or higher into the air and catches it in either hand. This is done while remaining in the lying position. Scoring is accomplished by counting the number of successful catches in ten trials.

3. **Ball-bounce Test.** Standing in the center of a 6-foot circle, the subject attempts to bounce a volleyball on the top end of a bat without leaving the circle. Scoring is done by noting the number of consecutive volleys up to ten. The total score is the sum of the scores on ten trials.

4. **Basketball Shooting Test.** This test is scored by counting the number of baskets made in twenty free throws from the foul line.

It may appear to the reader that the Adams test items are skill test items or items related to general motor ability. More research seems necessary in order to show the degree to which the tests appraise ability to learn motor skills.

Adams worked out a regression equation for college men, which employs the total score as follows:

[53] Arthur R. Adams, "A Test Construction Study of Sport-type Motor Educability Test for College Men," unpublished doctoral dissertation, Louisiana State University, Baton Rouge, La., 1954.

Sport educability score = 7.2 (wall volley) + 17.3 (tennis ball catch) + 2.7 (ball bounce) + 19.2 (foul shooting)

Motor Educability and Learning Scores

Instead of measuring motor intelligence or educability by having the subject respond to a certain test, it is possible to measure this factor by an analysis of *learning scores;* that is, scores that show a rate of learning. Henry, for example, set up three motor learning experiments involving improvement due to practice in the vertical jump, speed in performing an arm movement, and balancing.[54] Correlations were fairly high between initial and final degrees of skill (.63–.85). Raw learning scores were negatively correlated with initial skill but had no significant relation with final attainment. By holding statistically constant the influence of individual differences in initial skill, Henry computed a set of corrected learning scores that exhibited approximately zero relationship with initial skill. These derived scores were found to correlate fairly well with raw learning scores (.58) and with final scores (.61). They appear to offer, therefore, a valid measure of motor educability. More research is indicated in this technique.

MEASURES OF FLEXIBILITY

No study of motor ability would be complete without a discussion of flexibility. It is also conceivable that the topic might well be placed in the area of posture and body mechanics. In either case it is sufficient to say that human performance is related to the degree of tension that exists in the structure of the connective tissues and muscles.

In present-day society, man is subjected to more and more bombardment of stimuli. The impressions from sights and sounds, coupled with the numerous threats to life and limb and psychological stresses of anxiety, apprehension, worry, and fear, combine to keep the muscular system in a neurogenic state of contraction. It becomes difficult to relax even when one desires to. Improper relaxation, like obesity, is a disease of civilization. Children, especially overactive ones, do not escape it.

Under these circumstances, the tension syndrome is reflected in a reduced range of movement in some joints. Body flexibility per se is lowered. The generally smooth performance in body mechanics and the ability to execute motor skills efficiently appears to be somewhat affected. This is not a new viewpoint. In fact, Kraus and his associates included a test of back and leg flexibility in their test battery because so many persons who were just short of being emotionally disturbed showed evidence of hypokinetic disease in which tension limited their degree of movement. This is further supported by the historical works of Edmund Jacobsen and by the efforts of Josephine

[54] Franklin M. Henry, "Evaluation of Motor Learning When Performance Levels Are Heterogeneous," *The Research Quarterly,* 27:176–181, May, 1956.

Rathbone to associate the relaxed body with efficient physical performance, serenity, and peace of mind.

Kraus-Weber Floor-touch Test of Flexibility

As already pointed out in Chapter 7, the sixth test item in the Kraus-Weber Tests is the measure of back and leg flexibility. According to the authors, a person who cannot bend over and touch the floor with the fingertips without bending the knees exhibits less than a minimum level of physical fitness. In the widespread school use of the Kraus-Weber battery, Kraus's viewpoint was generally substantiated. Large numbers of children unable to pass the test of flexibility were likewise weak, substrength, and in general poor physical condition. To be completely objective, however, it should also be pointed out that large numbers of otherwise physically efficient school children failed to pass the flexibility test. Also, girls are substantially superior to boys in flexibility at all age levels, but this decreases with age.[55] Inflexibility increases in both sexes with age. Of course, there is the question of individual differences in joint contours, length of joint capsule and ligaments, and tightness of the overlying facia. Fox and Atwood questioned the Kraus-Weber test of flexibility on this basis.[56] So did Lawther in his analysis of the flexibility element in fitness, where he properly pointed out that flexibility must always be considered not as an end but as a means to an end.[57]

Certainly flexibility is related to physical condition. The elasticity difference between a one-year-old child and a seventy-four-year-old adult may be 80 per cent.[58] Some support for the Kraus-Weber test has come from Matthews, Shaw, and Bohnen, who found no significant relationship between the flexibility of the hip joint in the anteroposterior plane and the length of body segments.[59] In a later study into the relationship of trunk-plus-arm length (reach) to leg length in the ability to perform the toe-touch test, Broer and Galles found that reach and leg lengths were not important factors for persons with average body builds.[60]

[55] Marjorie Phillips et al., "Analysis of Results from the Kraus-Weber Test of Minimum Muscular Fitness in Children," *The Research Quarterly*, 26:314–323, October, 1955.

[56] Margaret G. Fox and Janet Atwood, "Results of Testing Iowa School Children for Health and Fitness," *Journal of Health, Physical Education and Recreation*, 26:20–21, September, 1955.

[57] John D. Lawther, "Flexibility for What," *Journal of Health, Physical Education and Recreation*, 27:23–24, March, 1956.

[58] Thomas K. Cureton, "Flexibility as an Aspect of Physical Fitness," *The Research Quarterly*, 12:381–390, May, 1941.

[59] Donald K. Matthews, Virginia Shaw, and Melva Bohnen, "Hip Flexibility of College Women as Related to Length of Body Segments," *The Research Quarterly*, 28:352–356, December, 1957.

[60] Marion R. Broer and Naomi R. G. Galles, "Importance of Relationship between Various Body Measurements in Performance of the Toe-touch Test," *The Research Quarterly*, 29:253–263, October, 1958.

Adapted Kraus-Weber Floor-touch Test

In order to increase objectivity Matthews measured flexibility of school children exactly as Kraus-Weber suggested, except for one change. The test is taken on a stall bar bench where a 20-inch scale is attached, so that the exact number of inches the subject reaches can be recorded. In this fashion the subject can be scored according to the number of inches reached above or below the toes rather than scored pass or fail. This makes for objectivity.

Wells Sit-and-Reach Test of Flexibility

Working with college women, Wells and Dillon devised a measure to improve upon the standing, bobbing type of test (where the subject, with arms hanging, bends forward and bobs four times before being measured on a vertical scale).[61] Wells built a graduated measuring scale that would rest on crossed stem boards raised 1 foot from the floor. The subject sits with feet firmly placed in footprints painted on the cross board. The arms are extended forward, with the hands placed palms down on the upper surface of the scale. After bobbing forward four times, the position of maximum reach is held on the fourth count. The most distant point reached on the fourth bob is held, and the score is the reading to the nearest half inch.

The test is highly reliable (.98). Moreover it correlates .95 with the Adapted Kraus-Weber Floor-touch Test.[62]

Protractor and Goniometer

Over the years joint flexibility has been measured with a transparent 180-degree protractor. When 12- to 15-inch arms are added to the protractor, one to remain stationary and follow the base line and the other movable, the device is referred to as a goniometer (see Figure 7-12).

Employing this type of measuring device it is possible to rate joint flexibility fairly well. Hall, for example, employed a form of measurement patterned after the early work of Cureton to appraise trunk flexibility forward and backward.[63]

1. **Trunk-forward Test.** The subject sits on the floor with legs extended, fingers laced behind the neck, and tries to touch the forehead against the ground. The knees are held down by a partner. By employing a vertical scale and noting the level of the eyebrows it is possible to measure the distance by which the subject misses touching the floor.

2. **Trunk-backward Test.** The subject lies face down on the ground with fingers laced behind the neck. The legs are held down by a partner by

[61] Katherine F. Wells and Evelyn K. Dillon, "The Sit and Reach: A Test of Back and Leg Flexibility," *The Research Quarterly*, 23:115–118, March, 1952.
[62] Donald K. Matthews, Virginia Shaw, and Melva Bohnen, *op. cit.*, p. 353.
[63] D. M. Hall, "Standardization of Flexibility Tests for 4-H Club Members," *The Research Quarterly*, 27:296–300, October, 1956.

grasping the ankles. The greatest height to which the chin can be raised is measured by a goniometer. The apex of the angle is placed at the large trochanter joint of the femur. The arc is measured through the center of the ear by a small wand attached to a cord held at tension by the coil spring on the protractor. The center of the ear instead of the chin is used as a landmark to avoid including the rotation of the head.

This same type of flexibility measure can be done more crudely by employing a vertical scale to note the height that the chin rises above the level of the floor.

A method employed to measure the elasticity of the hamstring tendons is one originally reported by Kraus and Weber where the subject is handled as follows:[64]

1. Have the subject lie on a table in a supine position completely relaxed.
2. Place the protractor with the base on the table next to the subject's hip joint.

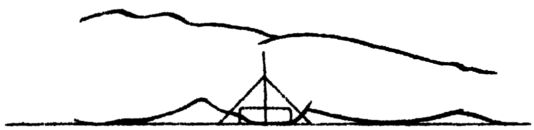


Fig. 11-5. Subject in Supine Position for Appraising Hamstring Flexibility.

3. Slide the hand gently under the subject's back and place the fingers lightly on the posterior spinal processes of the fourth and fifth lumbar vertebrae.
4. Have an assistant *slowly* raise both of the subject's legs holding them at the heels. The subject must remain perfectly relaxed during this procedure. The moment any motion is felt in the lumbar spine, it will indicate that the full length of the hamstrings has been reached; continued lifting of the legs will cause movement of the lumbar vertebrae in a posterior direction.

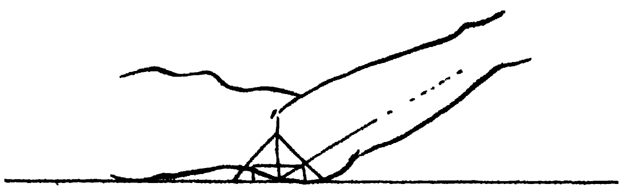


Fig. 11-6. Subject in Position for Appraising Hamstring Flexibility.

5. When this motion is felt, the assistant should stop lifting and hold the legs steady at the particular angle.
6. Adjust the arm of the protractor parallel to the length axis of the

[64] Hans Kraus and S. Weber, "Evaluation of Posture Based on Structural and Functional Measurements," *Physio-Therapy Review*, 25:6–12, November–December, 1945.

femur. Note the angle and record the number of degrees of the acute angle between the table level and the femur.

Leighton Flexometer

The range of motion in the joints has always been difficult to appraise accurately. In an effort to overcome this weakness, Leighton designed and built an instrument equipped with a rotating flat circular dial, graduated from 0 to 360 degrees, and with a movable pointer.[65] The instrument, called a *flexometer,* is ready to use when the direction of the dial and pointer, which always points upward, coincide. Both dial and pointer can be locked into position. To measure flexibility of a joint the instrument is strapped on the moving part to be tested. The dial is then locked at one extreme position, such as full flexion of the knee. The proper limb movement is made, such as full extension of the knee, and the pointer is locked at the extreme position. By noting the dial pointer the complete range of movement can be read in degrees. Leighton worked out twenty-one tests of joint flexibility employing the flexometer. Coefficients of reliability vary from .89 to .99 for these tests. Also, the range of motion for each test is presented in the form of averages in the cited reference.

SELECTED REFERENCES

Anderson, Theresa, and Charles H. McCloy: "The Measurement of Sports Ability in High School Girls," *The Research Quarterly,* 18:2-11, March, 1947.

Brace, David K.: "Studies in the Rate of Learning Gross Bodily Motor Skills," *The Research Quarterly,* 12:181-184, May, 1941.

Carpenter, Aileen, "Factors in Motor Educability," *The Research Quarterly,* 14:366-371, December, 1943.

———, "Studies in Motor Learning Gross Bodily Motor Skills," *The Research Quarterly,* 17:242-253, December, 1946.

Espenschade, Anne: "Development of Motor Coordination in Boys and Girls," *The Research Quarterly,* 18:30-43, March, 1947.

Hanley, Charles: "Physique and Reputation of Junior High School Boys," *Child Development,* 22:247-260, March, 1951.

Jones, Harold E.: "The Development of Physical Abilities," in National Society for the Study of Education, *Adolescence, 43d Yearbook,* Chicago, University of Chicago Press, 1944, chap. 6.

Kuhlen, Raymond G., and Beatrice J. Lee: "Personality Characteristics and Social Acceptability in Adolescence," *Journal of Educational Psychology,* 34:321-340, September, 1943.

Massey, Benjamin H., and Norman L. Chandet: "Effects of Systematic, Heavy Resistive Exercise on Range of Joint Movement in Young Male Adults," *The Research Quarterly,* 27:41-51, March, 1956.

[65] Jack R. Leighton, "A Simple Objective and Reliable Measure of Flexibility," *The Research Quarterly,* 13:205-216, May, 1942.

McCloy, Charles H.: "The Influence of Chronological Age on Motor Performance," *The Research Quarterly*, 6:61-63, March, 1935.

McCraw, L. W., and J. W. Tolbert: "Sociometric Status and Athletic Ability of Junior High School Boys," *The Research Quarterly*, 24:72-80, March, 1953.

Roloff, Louise L.: "Kinesthesis in Relation to the Learning of Selected Motor Skills," *The Research Quarterly*, 24:210-215, May, 1953.

Scott, M. Gladys, and Esther French: *Measurement and Evaluation in Physical Education*, Dubuque, Iowa, Wm. C. Brown Publishing Co., 1959.

Solley, William H.: "Relationship of Physical Development as a Factor in Motor Coordination of Boys Ages 10-14," *The Research Quarterly*, 28:295-304, October, 1957.

CHAPTER 12

Tests of Knowledge in Physical Education

In the ninth century B.C. Lycurgus, an astute creator of law in ancient Sparta, voiced this opinion when asked if Sparta should be enclosed by protective walls: "That city is well fortified which has a wall of men instead of brick." Lycurgus was not referring alone to strength of body but to strength of spirit as well.

Persons stand for the things they thoroughly understand. The wisdom to act in keeping with the aims and objectives of education is founded on knowledges and skills plus *understanding.* The virtues of being a physically educated person are most rewarding, but a person is not educated physically by mere development of motor skills and factual knowledge. One must understand why the skills are beneficial and why exercise is necessary and appreciate the role of recreational activity in the rich and full life. This does not come automatically; it has to be planned for during the years of youth.

The Measurement of Understanding

It is as necessary to administer written tests in physical education classes as it is to burn off energy and enjoy a game of skill. In fact, to be at all certain that a real understanding of physical education activities is being gained by pupils, knowledge and attitudes must be carefully appraised. Conceivably this can be more troublesome than measuring ability in isolated skills.

All too often appraisal of knowledge and understanding is overlooked in physical education classes. Yet if the cultural objective is to be reached and if a deep apprecation for the role of sports and physical activity in human life is to be obtained by the masses, more evaluation must be applied formally and informally. Cozens and Stumpf have eloquently stressed the need for sports education. They bring out clearly that "common interests, common loyalties, common enthusiasms . . . are the great integrating factors in any culture. In America, sports have provided this common denominator in as great a degree as any other single factor."[1]

Rice points out that somehow we need to educate future citizens how

[1] Frederick W. Cozens and Florence Stumpf, "American Sports from the Sidelines," *Journal of Health, Physical Education and Recreation,* 23:12–14, November, 1952.

to behave at games.[2] At far too many games, from elementary to professional circles, there is much undesirable spectator behavior. Much of this indicates an ignorance of game rules and procedures, as well as a lack of appreciation for the specialized performance and physical skills of the athlete. In short, education for sportsmanship requires more than a hit-or-miss program. It needs to be purposefully studied by pupils.

To be physically educated one does not necessarily have to possess skills in every major activity, but there should be some real understanding of the activities. If it is true that sports are a part of the great American tradition and the culture of western civilization, then the physical education teacher should evaluate the pupils, especially above the elementary school level, in such items as the following:

Sports terminology
Rules of the game
Strategy
Game equipment and playing areas
Athletic conditioning
Role of exercise in physical performance
Role of diet and sleep in physical performance
Historical background in special sports areas

An investigation in these areas will give some idea of the pupils' extended knowledge. This, coupled with skill performance appraisal, can provide a solid basis for grading. Moreover, such tests can serve as motivation devices by acquainting students with their level of ability. The Héwitt Comprehensive Tennis Knowledge Test is a good illustration of such a test. It deals with the history of tennis, equipment, rules of the game, playing situations, and fundamentals (see Appendix B).

Attitude Appraisal

As previously indicated it has never been easy to evaluate attitudes (see Chapter 5). Too many variables are involved in human behavior besides facts and skills. An effort, however, can be made to rate the attitudes of boys and girls toward physical activity in general. Do they like the games? Are they enthusiastic? Do they ask genuine questions? Do they take these activities home with them? What is their attitude at game time?

In an effort to measure the intensity of individual and group attitudes toward physical education as an activity course, Wear developed an inventory of 120 statements to be used with college students.[3] A short form of forty statements was also constructed. The subject has five choices in answering the question. He may "agree," "strongly agree," "disagree,"

[2] Sidney Rice, "Sports Fans Are a Menace," *Journal of Health, Physical Education and Recreation,* 23:14–16, May, 1952.
[3] Carlos L. Wear, "The Evaluation of Attitude toward Physical Education as an Activity Course," *The Research Quarterly,* 22:114–126, March, 1951.

"strongly disagree," or be "undecided." T-score tables, together with the statements, are included in the cited reference. Later, Wear developed two forms of a thirty-item attitude scale for use in physical education classes.[4] The scale has practical value and is worth studying.

McGee explored attitude appraisal by measuring the feelings of high school girls toward intensive athletic competition.[5] Here a seventy-item attitude scale was constructed to appraise the attitude of students, administrators, teachers, and parents. In this study the parents seemed to favor intensive athletic competition, whereas teachers and administrators were far less enthusiastic. This is a very complete scale for the topic, and scoring is similar to the Wear inventory.

What students of physical education think of themselves is sometimes worthwhile to know. A "who are you?" type of question was worked out by Bugental and Zelen.[6] It was later used to measure self-attitudes by Kuhn and McPartland with a high degree of predictiveness.[7] In this test twenty statements permit twenty answers to the simple question, "Who am I?" A time limit is established for the test to assure spontaneous, unstructured responses. Isenberger employed the test with college physical education personnel and found that the attitudes of women students differ considerably from those of the physical education teachers.[8] Later it was found that there is a relationship between physical education major students' self-attitudes and their motor ability.[9]

An attitude survey can be highly indicative of the true reaction of groups of students to the physical education program and its outcomes. Pupils can be asked to rate such statements as the following:

We have enough outdoor space for our games and activities.
We have enough time to shower and dress after class.
Our physical education teacher is friendly.
The physical education classes are very interesting.

A battery of statements such as these can be reliably scored and used by applying a five-interval scale such as "strongly agree," "agree," "undecided," "disagree," and "strongly disagree." Common language responses sometimes work well, too. Kelly, for example, checked the attitudes of students

[4] Carlos L. Wear, "Construction of Equivalent Forms of an Attitude Scale," *The Research Quarterly,* 26:113–119, March, 1955.

[5] Rosemary McGee, "Comparison of Attitudes toward Intensive Competition for High School Girls," *The Research Quarterly,* 27:60–73, March, 1956.

[6] James F. T. Bugental and Seymour L. Zelen, "Investigation into the Self-concept," *Journal of Personality,* 18:483–498, 1950.

[7] Manford Kuhn and Thomas S. McPartland, "An Empirical Investigation of Self-attitudes," *American Sociological Review,* 19:68–76, February, 1954.

[8] Wilma Isenberger, "Self-attitudes of Women Physical Education Major Students and of Women Physical Education Teachers," *The Research Quarterly,* 30:44–53, March, 1959.

[9] Wilma Isenberger, "Self-attitudes of Women Physical Education Major Students as Related to Measures of Interest and Success," *The Research Quarterly,* 30:167–178, May, 1959.

about space allotment, program, class organization, teacher effectiveness, and physical education values and used five responses: "I'm very sure," "I think so," "I don't know," "I doubt it," and "I disagree strongly."[10]

Test Preparation

Tests should not be given too often or be overly time-consuming. Pupils dress for the gymnasium or the playground and are expecting vigorous physical activity. The popularity of a program, especially in the lower grades, can be drastically reduced if too many written tests are used. Short ten- or twelve-minute quizzes may be given in connection with a unit of work. Students will keep up with what is being taught and will readily do outside reading when they know that short tests will frequently be given. In planning the test it is advisable to give wide coverage to the unit of work. It should reflect the course emphasis, consider terminology, spectatorship, facility layout, etc., as well as the more common items. A simple scoring procedure is recommended. Eventually it should be evaluated to see how well it appraises understanding. For a more elaborate discussion of test preparation, refer to Chapters 3 and 5.

A number of physical education knowledge and understanding tests have been developed and standardized. Several of these are worth looking over for local use. The cited references often contain the specific questions and occasional norm tables. Some of these tests are referred to below.

COMBINED KNOWLEDGE TESTS

Several investigators have prepared a number of knowledge tests and grouped them into one battery.

Stradtman and Cureton prepared a *Physical Fitness Knowledge Test* for secondary school boys and girls.[11] It was carefully prepared in the subject matter area of physical fitness, with special emphasis placed on the selection of vocabulary.

French, at the State University of Iowa, combined the activities of badminton, basketball, body mechanics, canoeing, field hockey, folk dancing, golf, bowling, deck tennis, handball, shuffleboard, tennis, tetherball, rhythms, soccer, softball, swimming, tumbling, track and field, and volleyball into one comprehensive knowledge test for college women.[12] Both a long form and a short form of the test are available. Norms have been constructed for the short form. Test reliability ranges from .70 to .88 for the long form and from .61 to .87 for the short form. The test is useful in measuring a student's advancement when given before and after a unit of instruction.

[10] Ellen Kelly, "Attitudes Are Important," *The Physical Educator,* 12:86–88, October, 1955.

[11] Alan D. Stradtman and T. K. Cureton, "A Physical Fitness Knowledge Test for Secondary School Boys and Girls," *The Research Quarterly,* 21:53–57, March, 1950.

[12] Esther French, "The Construction of Knowledge Tests in Professional Courses in Physical Education," *The Research Quarterly,* 14:406–413, December, 1943.

In a study of Women's College of the University of North Carolina, Hennis constructed several knowledge tests suitable for use in the college women's instructional program.[13] These are multiple-choice tests covering badminton, basketball, bowling, field hockey, softball, tennis, and volleyball. The tests are available from the author to be used as is or to serve as a guide in the construction of teacher-made tests. A wide sample of colleges was taken before the test was completely validated. Reliability is satisfactory (.72 to .81), and percentile norms have been prepared.

Scott, employing the index of discrimination and the difficulty rating, carefully prepared knowledge examinations in badminton, swimming, and tennis.[14] They were built for general teacher use among college women through the committee efforts of the Central Association of Physical Education for College Women. The battery of multiple-choice and true-false statements has been used in a number of institutions quite successfully. Reliability varies from .72 to .79 in badminton to .89 for the elementary swimming knowledge test. Scott has also developed a grading system for the badminton test based upon the test scores of 350 badminton players.

A very good set of written tests for use with college physical education classes for men were developed and are in use at the University of Florida. Men taking required physical education are tested periodically in the following sports:

Beginning golf	Advanced golf
Beginning handball	Advanced handball
Beginning tennis	Advanced tennis
Swimming	Touch football
Basketball	Volleyball
Softball	

The raw scores on these tests are related to T-score norm tables, and the final score of the subject's knowledge and understanding becomes a part of the semester grade. Sample mimeographed copies of these tests and scoring tables can be obtained from The Dean, College of Physical Education and Health, University of Florida, Gainesville, Florida.

BADMINTON

In addition to the efforts of Scott and others to appraise badminton knowledge, there have been other tests equally as satisfactory. Phillips, for example, subjected 1,471 college women to a test of fifty-five true-false

[13] Gail M. Hennis, "Construction of Knowledge Tests in Selected Physical Education Activities for College Women," *The Research Quarterly,* 27:301–309, October, 1956.

[14] See M. Gladys Scott, "Achievement Examinations for Elementary and Intermediate Swimming Classes," *The Research Quarterly,* 11:100–104, May, 1940; "Achievement Examinations for Elementary and Intermediate Tennis Classes," *The Research Quarterly,* 12:40–45, March, 1941; "Achievement Examinations in Badminton," *The Research Quarterly,* 12:242–248, May, 1941.

multiple-choice questions.[15] The students were grouped into two categories: homogeneous and heterogeneous. Reliability of the test was .87 and .92, respectively. A mean difficulty rating for the total test on all subjects was 50.8 per cent. Both T scales and percentile scales are available to rate beginning and intermediate badminton players.

At the University of Washington, a committee of women under Katherine Fox constructed a badminton knowledge test for beginners, which consisted in 106 multiple-choice, true-false, and identification questions.[16] Test reliability, as the result of the split-halves technique and application of the Spearman-Brown prophecy formula, is .90. Content validity and item validity is also satisfactory. The total list of test questions plus a scoring table appear in the cited reference. An analysis of the questions reveals a rather complete coverage of badminton rules, techniques, strategy, and special terminology.

GOLF

Of the more recent efforts to appraise understanding of golf, the test developed by Waglow and Rehling has considerable merit.[17] This 100-item true-false test is used in the required physical education program for college men and was validated by an analysis of well-known books on golf. Employing the split-half method and the Spearman-Brown prophecy formula the authors determined test reliability at .82. Copies of the test questions and T-score standards can be obtained from the primary author at the University of Florida, Gainesville, Florida, or from the cited reference.

It should not be difficult to build a set of golf knowledge questions for use with secondary school boys and girls. It is suggested that a trial set be put together and checked for difficulty rating and item discrimination. The satisfactory questions could then be retained and used to build a T-score or percentile table for local use.

GYMNASTICS

An experimental test of apparatus gymnastics was administered by Gershon, of Wisconsin State College, to 586 male professional physical education students at twenty-one colleges and universities.[18] After item analysis a revised test of 100 items was produced. This test in its final form

[15] Marjorie Phillips, "Standardization of a Badminton Knowledge Test for College Women," *The Research Quarterly,* 17:48–54, March, 1946.

[16] Katherine Fox, "Beginning Badminton Written Examination," *The Research Quarterly,* 24:135–142, May, 1953.

[17] I. F. Waglow and E. H. Rehling, "A Golf Knowledge Test," *The Research Quarterly,* 24:463–470, December, 1953.

[18] Ernest Gershon, "Apparatus Gymnastics Knowledge Test for College Men in Professional Physical Education," *The Research Quarterly,* 28:332–341, December, 1957.

was administered to 940 college men in professional physical education, enrolled in forty colleges and universities across the United States. This permitted Gershon to develop a set of national norms in the form of T scales. Moreover, a five-point evaluation schedule was formed to be used as a guide for grading students. Reliability, objectivity, and essential validity have been determined, and the subject matter was evaluated by a competent jury of educators. All test items proved significant for use at the 5 per cent level of confidence or better. Copies of the test, with directions, scoring key, and evaluation schedules can be obtained by any qualified physical education instructor from the author at Wisconsin State College, La Crosse, Wisconsin.

FIELD HOCKEY

Some years ago Deitz and Freck developed a test of field hockey knowledge, but the basis for its validity was not determined.[19] With this as a background Kelly and Brown built an objective eighty-eight-item multiple-response test designed to appraise understanding of rules, techniques, coaching procedures, and game officiating.[20] The measure warrants use, for it has satisfactory validity and reliability and has had close analysis in application.

HANDBALL

Standardized handball tests are difficult to find. Local tests are available in some institutions where handball is taught as part of the required physical education curriculum. As already mentioned, beginning and advanced tests are used regularly at the University of Florida and are complete with appropriate scoring scales. Other knowledge tests have been used at Yale University and at the State University of New York, Oswego, New York. The instructor of handball is encouraged to build his own test questions at the local level and apply the techniques of test construction already referred to.

ICE HOCKEY

In view of the increase in the number of young persons and adults following ice hockey from Minnesota to Colorado and from New Hampshire to Texas, it seems advisable to appraise student understanding in this area of sport. Actually very little test construction has been done, even experimentally. Several sample test questions of the true-false type were de-

[19] Dorothea Deitz and Beryl Freck, "Hockey Knowledge Tests for Girls," *The Research Quarterly,* 11:366–372, June, 1940.

[20] Ellen D. Kelly and Jane E. Brown, "The Construction of a Field Hockey Test for Women Physical Education Majors," *The Research Quarterly,* 23:322–329, October, 1952.

veloped by Willgoose.[21] Although they were not thoroughly appraised for discriminatory powers, they are useful as a basis for further experimentation.

An attempt was made several years ago to construct an ice hockey test for girls. Brown developed a short, eighteen-item true-false test in accordance with girls' ice hockey rules.[22] As no attempt was made to validate it and because rules change in time, the test may have limited use at present.

SOCCER

Written tests of knowledge usually complement skill tests in sport. With this in mind Heath and Rodgers built a 100-item true-false test for use with fifth- and sixth-grade boys.[23] Test validity is difficult to comprehend, but reliability was determined to be .90 with the chance-halves and Spearman-Brown formula techniques. T scales are provided in the cited reference. Here again is a sport weak in standardized tests of knowledge. It is not uncommon, however, to find nonstandardized tests in use in elementary and secondary school physical education classes.

SOFTBALL

This is another sport in which good written tests are lacking. As with their other skill tests in the early 1930s, Heath and Rodgers constructed a reliable softball test for fifth- and sixth-grade boys, complete with T-scale tables and based on 100 true-false statements.[24] Some of the questions may now need revision.

One of the most thoroughly developed written tests in physical education was constructed at the University of Florida, by Waglow and Stephens, to determine the extent of softball knowledge of college students.[25] This test has been widely applied in Florida and elsewhere. Reliability is .83 based on the odd-even technique and Spearman-Brown prophecy formula. Validity of the test questions was appraised by difficulty rating and index of discrimination. In fact, each test question as given in the cited reference has the difficulty rating and index of discrimination recorded next to it so the instructor using the test can personally appraise the question. There are sixty true-false questions, twenty-five fill-in items, five fair-or-foul questions, and

[21] See chapter on ice hockey in John H. Shaw et al., *Selected Team Sports for Men,* Philadelphia, W. B. Saunders Company, 1952, pp. 65–66.

[22] Harriett M. Brown, "The Game of Ice Hockey," *Journal of Health and Physical Education,* 6:28–29, January, 1935.

[23] Marjorie L. Heath and Elizabeth G. Rodgers, "A Study in the Use of Knowledge and Skill Tests in Soccer," *The Research Quarterly,* 3:33–38, December, 1932.

[24] Elizabeth G. Rodgers and Marjorie L. Heath, "An Experiment in the Use of Knowledge and Skill Tests in Playground Baseball," *The Research Quarterly,* 2:113–116, December, 1931.

[25] I. F. Waglow and Foy Stephens, "A Softball Knowledge Test," *The Research Quarterly,* 26:234–243, May, 1955.

ten plus-or-minus questions. They cover most phases of play, rules, equipment, and history. T scores for the softball knowledge test are available for the one hundred items. Also, for grading purposes letter marks have been arranged. The cited reference has all the information necessary for giving the test.

SWIMMING

The American Red Cross, the Boy Scouts of America, and the Young Men's Christian Association have for years required an understanding of a fairly large body of knowledge in connection with swimming and related aquatic activities. Numerous local tests have been used in connection with skill tests to determine swimming competencies.

A written examination suitable for college swimming classes was developed by the Central Association of Physical Education for College Women, in which fifty-six multiple-choice and true-false items are given at the elementary level.[26] A total of fifty-eight tests appear on the battery for the intermediate swimming level. Test reliability for both examinations is reported above .87 on the basis of 100 cases each. The complete examination is reported in the cited reference.

TENNIS

A tennis knowledge test for women, which can be used to improve grading practices, was worked out at the University of Washington, in the Women's Physical Education Department.[27] During the experimental phase a 100-item test was given eighty-seven students; it covered tennis fundamentals, advanced strokes, footwork, timing, and position. It also covered court position and game strategy, history, equipment, and court markings. Reliability is .84 according to the split-halves method. Curriculum validity was ascertained by a review of textbooks and knowledge tests already accepted. Care was taken to relate test items to the curriculm content, and item discrimination was checked in order to know which questions to omit. When these details were completed, the revised examination of 128 questions was tried out on 343 beginning and intermediate students. The test seems to be worth using and requires approximately fifty minutes to administer. A complete listing of all 128 test questions is printed in the cited reference.

As already indicated, the Hewitt Test of Tennis Knowledge is a very good one embracing a wide coverage of the area.[28] See Appendix B. Here is

[26] M. Gladys Scott, "Achievement Examinations for Elementary and Intermediate Swimming Classes," *The Research Quarterly,* 11:100–111, May, 1940.

[27] Marion Broer and Donna Mae Miller, "Achievement Test for Beginning and Intermediate Tennis," *The Research Quarterly,* 21:303–313, October, 1950.

[28] Jack E. Hewitt, "Comprehesive Tennis Knowledge Test," *The Research Quarterly,* 8:74–80, October, 1937.

an examination with a reliability coefficient of .95, which correlated positively, but low, with Dyer's test. Scoring key and tables are available from the author at the Department of Physical Education, University of California, Riverside, California.

VOLLEYBALL

Using volleyball teachers as judges, Langston verified the curricular validity of his 100-item volleyball examination.[29] Item analysis was used to find out how well the questions discriminated. The seventy true-false and thirty multiple-choice statements appraise understanding of volleyball history, set-up, spike, passing, net recovery, service, block, offensive and defensive strategy, game rules, and officiating. Examinations, together with a scoring key, can be obtained from Langston at Eastern New Mexico University, Portales, New Mexico.

SELECTED REFERENCES

Green, H. A., and A. N. Jorgensen: *Use and Interpretation of High School Tests,* New York, Longmans, Green & Co., Inc., 1936, chap. VIII.

Marcuse, F. L.: "Objective and Subjective Examinations," *School and Society,* 72:222–223, Oct. 28, 1950.

Murphy, Mary A.: "Grading Student Achievement in Golf Knowledge," *The Research Quarterly,* 5:83–87, March, 1934.

Ross, Clay C.: *Measurement in Today's Schools,* 3d ed., Englewood Cliffs, N.J., Prentice-Hall, Inc., chaps. 4 and 5.

Seaton, Don C., Irene A. Clayton, Howard C. Leibee, and Lloyd Messersmith: *Physical Education Handbook,* 3d ed., Englewood Cliffs, N.J., Prentice-Hall, Inc., 1959, pp. 277–343.

Shaw, John H., Carl A. Troester, and Milton A. Gabrielson: *Individual Sports for Men,* Philadelphia, W. B. Saunders Company, 1950; see end of each chapter.

[29] Dewey F. Langston, "Standardization of a Volleyball Knowledge Test for College Men Physical Education Majors," *The Research Quarterly,* 26:60–66, March, 1955.

CHAPTER 13

The Measurement of Body Physique

Let me have men about me that are fat;
Sleek-headed men and such as sleep o' nights.
Yond Cassius has a lean and hungry look;
He thinks too much: such men are dangerous.
William Shakespeare, Julius Caesar, i, 2

From the time of Hippocrates, who urged his students to study the physique of the patient, to that of Draper and his studies in constitutional medicine, man has been aware of a relationship between body build and overt behavior. Literature, with its lengthy cast of characters, abounds in temperaments and actions in keeping with physical appearance and physique. Dickens could not paint a picture of old Scrooge as a jolly roly-poly person any more than Shakespeare could picture Cassius that way.

None of this is strange, for man behaves as he does because of what he is. Structure precedes function. Behavior, observed or unobserved, is related to physical constitution and it is modified by environment. The skyscraper stands up in strong winds, rain, snow, summer heat, and internal wear and tear from use. Its durability is directly related to the quantity and quality of steel, iron, cement, and wood in its structure. Its total morphology and the very tolerance of its "tissues" relate to its unique function. The same can be said of the human body. The nature and quality of its structure determines in large part its mode of survival.

Constitutional Psychology, Health, and Physical Education

Because mind and body are a single unit and because there are individual differences in structure and in function, a study of body build related to total behavior is important to full understanding of health status and physical education.

Constitutional psychology is the study of the psychological aspects of human behavior as they are part of the morphology and physiology of the body. The focus is on the whole person. In short, it stresses and classifies individual differences, not to magnify or isolate them, but to see better how they relate to total behavior. This permits the health educator to inquire into

the health characteristics of various morphological types and into the sensitivity of personality. It encourages the physical education teacher to analyze body build carefully for physical fitness, athletic behavior, and general motor characteristics. The classification of pupils according to constitutional type can provide a basic framework for understanding pupil attitudes, interests, and general adaptability to physical exercise and social contacts in the gymnasium-playground laboratory.

The educational implications of constitutional psychology are most impressive. It is readily apparent that the more means the individual teacher has for observing and appraising behavior, the more likely he or she is to reach specific educational goals. The significant value of constitutional psychology comes when educators refrain from interpreting the work as an exact science and accept it as an indication of the tendencies various persons have toward certain kinds of behavior.[1] The key word here is *tendencies.* Certainly knowledge of a person's constitutional make-up is not the same as complete knowledge of character. Persons with about the same temperamental formula may behave in different ways and exhibit very different characters. In short, temperamentally similar persons can make dissimilar uses of their constitutional endowments. This is where environmental influences are brought to bear.

It would appear, when the educational needs of children are being considered, that those of differing constitutional make-up respond differently to educational influences. In all probability, if we don't know how to treat children as we think we should, it's because we have not learned to tell them apart. *When we can tell them apart through some common frame of reference, we will more thoroughly understand human behavior.*

Historical Background

Hippocrates designated two fundamental physical types: the *phthisic habitus* and the *apoplectic habitus.* The former was long and thin and particularly susceptible to tuberculosis (phthisis); the latter was heavy and solid, with a predisposition to diseases of the circulatory system leading to apoplexy. From that time on numerous philosophers, medical researchers, and anthropologists attempted to create a pattern or frame of reference classifying and using body build.

Following the early work of Halle (1797), two French anatomists, Gall and Spurzheim, originated a classification of *type digestif, type musculaire,* and *type cérébral.* This terminology was used by Rostan in Paris, in 1828. With the advent of the scientific spirit and method, correlation, measurement, and classification became known to research. Anthropometry, the physical measurement of man, began to blossom as a field of scientific endeavor. The Italian anthropologist, di Giovanni, set up three combinations

[1] Carl E. Willgoose, "Educational Implications of Constitutional Psychology," *Education,* 73:1–8, December, 1952.

of body build (1885). His loyal pupil, Viola, developed a morphological index, involving eight trunk-abdominal measurements and the length of one arm and one leg.[2] Then followed a pupil of Viola, named Sante Naccarati, who showed a low positive relationship between intelligence and Viola's index. Of greater promise were his studies in morphology and human temperament.[3] (Temperament as used in this text means, generally, the level of personality just above the physiological function and below acquired attitudes and beliefs. It is a level of behavior where man expresses his desires and motivations by interacting dynamically with others.)

During the above period considerable research was being carried on by anthropologists and medical physical educators, such as Hitchcock, Sargent, and Savage, on the relationship between all types of body measurements involving height, weight, cephalic index, bi-iliac diameter, subcutaneous fat, shoulder width to hip width, chest diameter, etc For a review of some of this history, see Chapters 1 and 7; see also Chapter 4, which refers to the ACH Index, Pryor's Tables, Montessori, Meredith Height-Weight Tables, and the Wetzel Grid.

Taking the lead from Montessori (1913), Bryant and Goldthwait, in 1915, proposed a three-category classification of carnivorous (linear), normal, and herbivorous (stocky) types of body build. Several years later Stockard wrote of the linear, intermediate, and lateral types.[4] These investigators were not the only ones to think in terms of three types of human constitution. Kretschmer (1926), working with mental patients, observed that most of his circular insanity cases (manic-depressives) appeared to belong in the macrosplanchnic or large viscera group, whereas the schizophrenes (split mind) were the microsplanchnic or small viscera types.[5] From this Kretschmer employed the Greek terms *pyknic* and *asthenic* to represent two extremes of structure: compact, and slight and without strength, respectively. To represent the vigorous type of body build falling in between these extremes, he then added the *athletic type*. This work was supported by Mohr and Gundlach, who examined 254 prison inmates at Joliet, Illinois.[6] Campbell, in 1932, and Garvey, a year later, used systems similar to that of Kretschmer to study the relationship between physique and mental disease.[7,8]

[2] A. di Giovanni, *Clinical Commentaries Deduced from the Morphology of the Human Body*, 2d ed., translated by J. J. Eyre, London and New York, 1919.

[3] Sante Naccarati and H. E. Garrett, "The Relationship of Morphology to Temperament," *Journal of Abnormal Social Psychology*, 1911, 1924.

[4] Charles R. Stockard, "Human Types and Growth Reactions," *American Journal of Anatomy*, 31:261-288, 1922–1923.

[5] E. Kretschmer, *Physique and Character*, 2d ed., translated by W. J. H. Sprott, New York, Harcourt, Brace and Company, 1926.

[6] G. H. Mohr and R. H. Gundlach, "The Relation between Physique and Performance," *Journal of Experimental Psychology*, 10:117–157, April, 1927.

[7] K. Campbell, "Relation of the Type of Physique to the Types of Mental Diseases," *Journal of Abnormal Social Psychology*, 27:147–151, September, 1933.

[8] C. R. Garvey, "Comparative Body Builds of Manic-depressives and Schizoid Patients," *Psychological Bulletins*, 30:567–568, October, 1933.

The work of Davenport, in 1923, and Draper, in 1930, threw much favorable light on constitutional research. Draper classified persons according to the general proportions of those he discovered to be particularly predisposed to certain diseases.[9] His *tubercular type* and *gastric ulcer type* are both slender or linear types, somewhat like Hippocrates' phthisic habitus. The *gall bladder type* and the *pernicious anemia types*, on the other extreme, are both laterals, or of stockier builds.[10]

Kretschmer's work in the area of cycloid and schizoid temperament is valuable basic research relating constitution to clinical behavior; but his body types were too loosely classified. Like many in the past, they made possible a comparison of structure to function without being scientifically specific. There was no effort to *structure* his three types or to scale the relative strengths and weaknesses of each type.

In the early 1920s Sheldon became impressed with the work of Naccarati at Columbia University and the findings of Kretschmer in Germany. As a medical doctor with an additional degree in psychology and philosophy, Sheldon had an exceptionally fine background to appreciate the interrelationship of mind and body, of psyche and soma. Sheldon set out on a long-range investigation to make a completely methodical approach to the study of human physique—a morphological typology.[11] The succeeding work, extending over a thirty-five-year period, is unique. It is incomplete, and yet it offers great immediate as well as potential value to health and physical educators. Already Sheldon's work has been related to medicine, aviation, education, sociology, and industry. A most enlightening study relates it to human values and happiness.[12] His work on the association between body types and delinquent behavior is a milestone in case-study techniques.[13]

The Somatotype Defined

In order to look over a large number of physiques at one time, Sheldon took full-length photographs of 4,000 male college students in a standardized posture from the front, side, and rear positions on a single film. By assembling the photographs in series it was possible to observe a definite orderliness of structure through anthroposcopy (with the unaided eye). After defining the dimensions of variation, the body types were classified according to three dominant morphological components. These components in combination

[9] G. Draper, *Disease and the Man*, London, Routledge & Kegan Paul, Ltd., 1930.

[10] For an excellent discussion of early body measurement designed to appraise physical status, see Charles H. McCloy, *Appraising Physical Status: The Selection of Measurement*, Iowa City, Iowa, State University of Iowa, 1936.

[11] W. H. Sheldon, S. S. Stevens, and W. B. Tucker, *The Varieties of Human Physique*, New York, Harper & Brothers, 1940.

[12] See especially Charles William Morris, *The Open Self*, Englewood Cliffs, N.J., Prentice-Hall, Inc., 1948; *Varieties of Human Value*, Chicago, University of Chicago Press, 1956.

[13] William H. Sheldon et al., *The Varieties of Delinquent Youth*, New York, Harper & Brothers, 1949.

constitute the somatotype, a word derived from the Greek *sōma*, or body.[14] Working at Harvard University, Columbia University College of Physicians and Surgeons, the Oregon Medical School, and twenty-six hospitals and medical institutions over a period of many years, Sheldon has built his findings on 46,000 men, aged eighteen to sixty-five, and several thousand women. The somatotype is a morphological (structural) frame of reference. It is a biological identification tag, which at the motivational level is expressed in temperament.[15]

By relating the development of the structure to the embryonic stages of growth, that is, endoderm, mesoderm, and ectoderm, Sheldon postulated three components of body build: endomorphy, mesomorphy, and ectomorphy.

1. Endomorphy. In endomorphy the digestive system dominates the body economy. A prominence of roundness of the various regions of the body is evident, with much concentration in the center. The head is large and round, a thick short neck, broad thick chest with fatty breasts, short arms, a large abdomen full above the navel, heavy fat buttocks, generally short heavy legs, narrow shoulders, and broad hips, with fleshy area above the crest of the ilium. In Thurber's words, this is the pear-shaped person, whose somatic structures are relatively undeveloped and exhibit softness.

This component is associated with slow reaction and relaxation in posture and movement.

2. Mesomorphy. A mesomorphic person is distinguished by his heavy, hard, rectangular outline. He possesses large bones, which are very prominent. He is unusually well muscled. He has very prominent facial bones, a long strong neck, fairly low thoracic trunk, broad shoulders with prominent clavicles, muscular upper arms, and very strong muscular forearms. The hands and fingers are usually large. The abdomen is heavily muscled with a slender low waist and narrow hips. He has heavy buttocks, strong, powerful legs, and a structure that is essentially hard, firm, and relatively tough. Uprightness and sturdiness prevail.

3. Ectomorphy. In ectomorphy, linearity and fragility prevail. The ectomorph has a slender, frail body structure with small bones and thin segments, a relatively large head with bulbous forehead, small facial bones, sharp nose, pointed chin, long slender neck, long narrow chest, tendency toward winged scapulas and round shoulders, long arms, and a very flat abdomen with the hollow above the navel. The buttocks are inconspicuous, the legs are long and thin with pipestem bones, and general musculature is not marked. In fact, the skin seems merely stretched over the bones with little musculature showing.

[14] See W. H. Sheldon, *Atlas of Men*, New York, Harper & Brothers, 1954 (*Atlas of Women* to follow when completed).
[15] William H. Sheldon and S. S. Stevens, *The Varieties of Temperament*, New York, Harper & Brothers, 1942, chap. 3.

This component is associated with overly fast reactions and hypersensitivity. There is marked postural restraint and movement tightness.

It is obvious that all persons possess in varying degree each of the three components. Sheldon used a three-dimensional classification, with which he developed the proper frame of reference to grade each component numerically from 1 to 7 on a 7-point scale. For this purpose a minimum of seventeen diameter measurements was carried out by detailed examination of the photographic negatives. These time-consuming appraisals were made in five regions of the body. The head, face, and neck constituted the first region; the second region was the thoracic trunk, the third the arms, shoulders, and hands, the fourth the abdominal trunk, and the fifth the legs and feet. It was determined early in the research that measurements from a clear negative are far more reliable than those taken directly on the body of the subject.

In order to save time in determining the somatotype from a series of measurements, a machine was built into which the measurement data could be entered. By manipulating switches the correct somatotype was disclosed. Since then, however, the IBM computer has been programmed with accurately designed and stored instructions and the appropriately devised electronic wiring pattern so it can be fed punched data cards. It then performs in a matter of seconds the operations of comparison, elimination, and selection and arrives at the somatotype.

Determining the Somatotype

The method of scaling the three components from 1 to 7 is relatively easy. In each instance, the first numeral in the sequence refers to endomorphy, the second to mesomorphy, and the third to ectomorphy. Thus, a person rated 711 would be an extreme endomorph. The first digit, 7, is the top of the endomorphy scale, whereas the next two digits show only the slightest trace of mesomorphy and ectomorphy, respectively.

A 171 is an extreme mesomorph. An extreme ectomorph is a 117. A balanced build would be classified 334, 343, 443, 444, or 333. These basic differences in physical structure, classified by the three components, represent the basic somatotype.

Although a person is rated according to the three basic components, most physiques are dominated by only two. The practice of naming a type according to the strongest component and applying the second strongest component as the descriptive adjective has been generally accepted. Thus a person rated 641 would be primarily an endomorph, with a secondary component of mesomorphy; he would be classified as a mesomorphic endomorph. A 461 would be classified as an endomorphic mesomorph and a 146 as a mesomorphic ectomorph.[16]

[16] It is proper practice to employ halves in rating a component. A somatotype 5½41 would be slightly stronger in the first component than one rated 541.

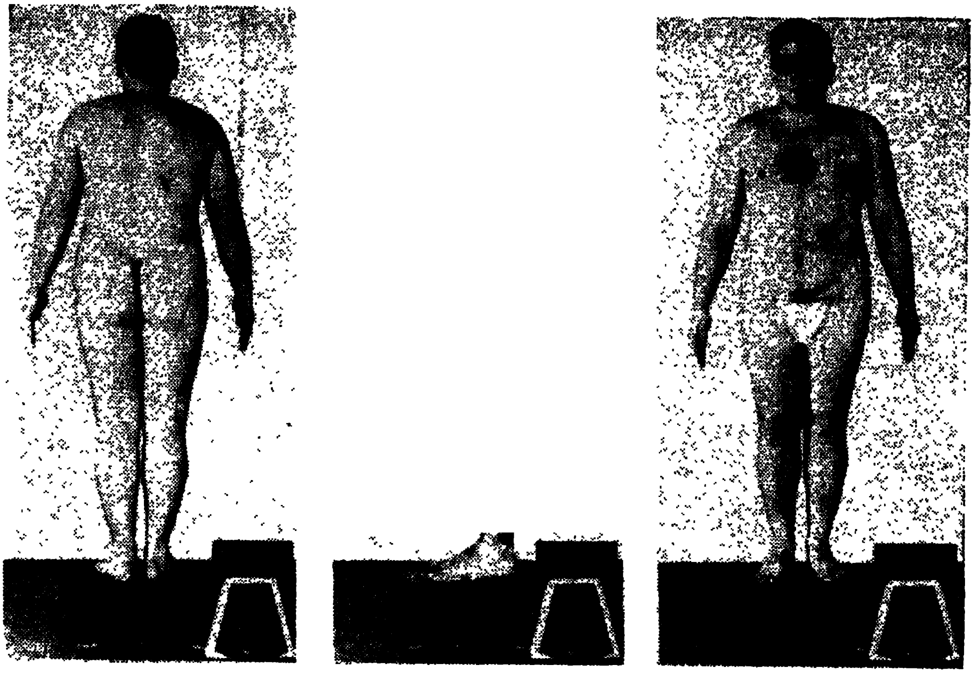


Fig. 13-1. Mesomorphic Endomorph (Somatotype 641).

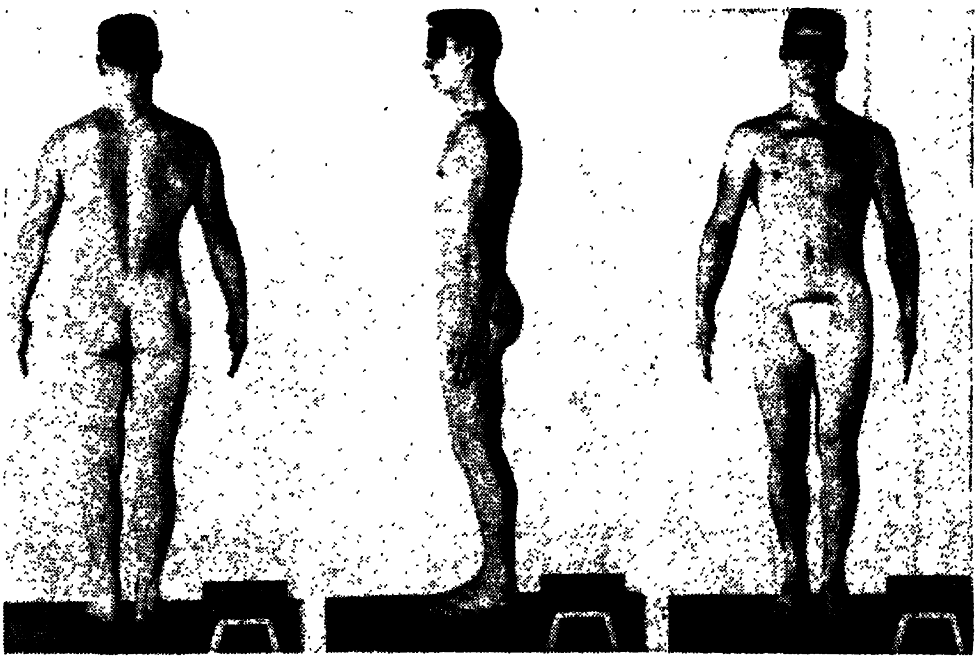


Fig. 13-2. Extreme Mesomorphy (Somatotype $1^{2}6^{2}1$).

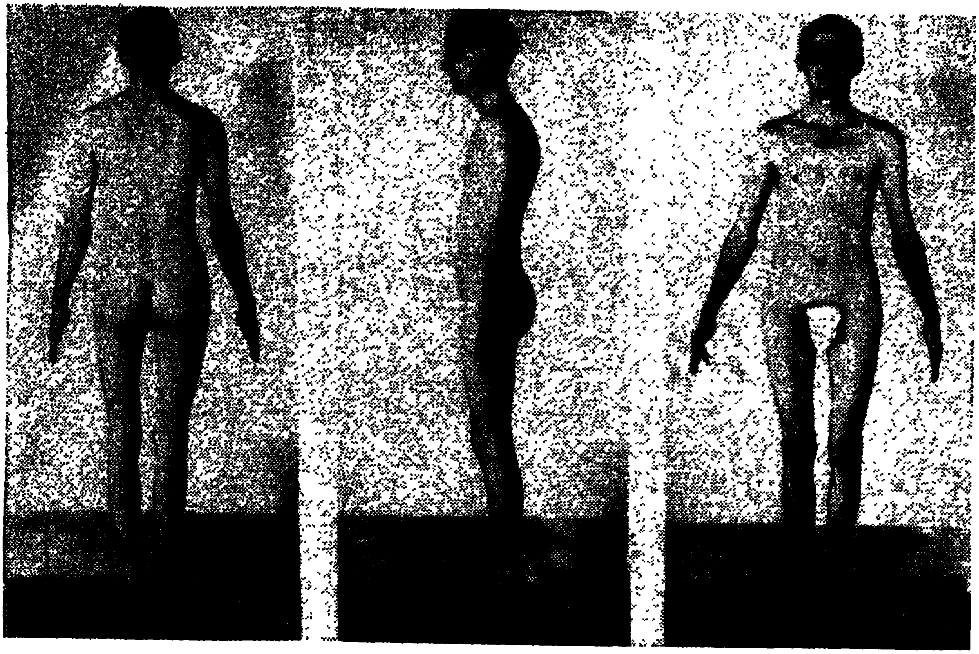


Fig. 13-3. Extreme Ectomorphy (Somatotype $1^2 1^2 6^2$)

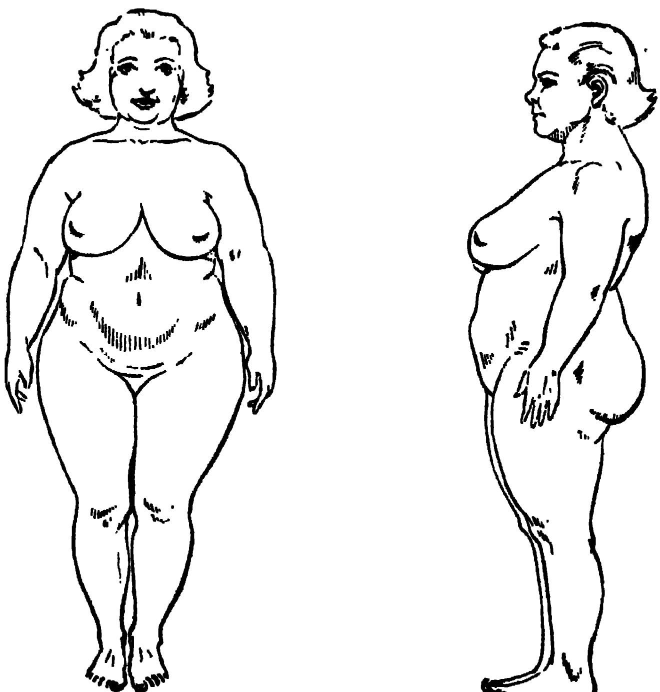


Fig. 13-4. Classic Female Somatotype 711.

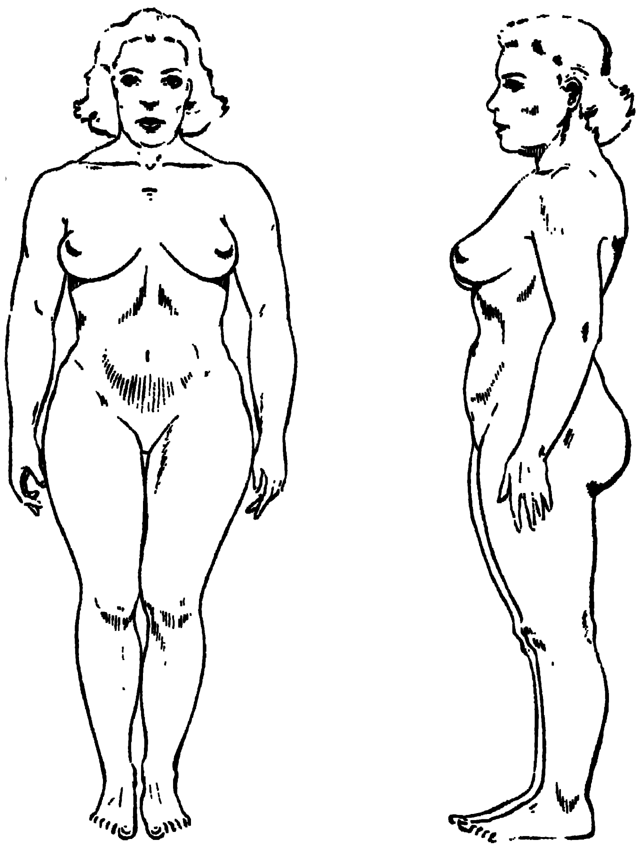


Fig. 13-5. Classic Female Somatotype 171.

When not speaking of somatotypes by the three-digit rating, it is proper to use descriptive terms. The classifications follow:

Extreme endomorph	(5^2, 6, 6^2, 7) 621, 712
Mesomorphic endomorph	642, 543, 532
Endomorph-mesomorph	443, 552, 4^24^22
Endomorphic mesomorph	462, 453, 35^22
Extreme mesomorph	(5^2, 6, 6^2, 7) 271, 16^22
Ectomorphic mesomorph	264, 354, 25^23
Extreme Ectomorph	(5^2, 6, 6^2, 7) 127, 216
Mesomorphic ectomorph	146, 245^2, 23^26
Mesomorph-ectomorph	144, 155, 244
Balanced (equal components)	$3^23^23^2$, 444
Endomorphic ectomorph	415, 325, 425
Endomorph-ectomorph	343, 434, 424
Ectomorphic endomorph	514, 613, 524
Endopene*	145, 13^26, 136
Mesopene	424, 4^225, 415
Ectopene	642, 541, 5^23^22

* Endopenes, mesopenes, and ectopenes are very low in their respective components. The mesopene, for example, has dominance of the first and third components, such as 424, 425, 415. An endomorphic ectomorph is a mesopene.

It should be stated at this time that the somatotype does not change essentially over the years, except to add or lose weight. Obviously it is the endomorphs that put on the most weight, with the mesomorphs adding some in the later years. Sheldon discovered early in his studies that an exact weight history of the subject is helpful. The ponderal index (height over the cube root of weight), when applied to somatotyping tables, places the

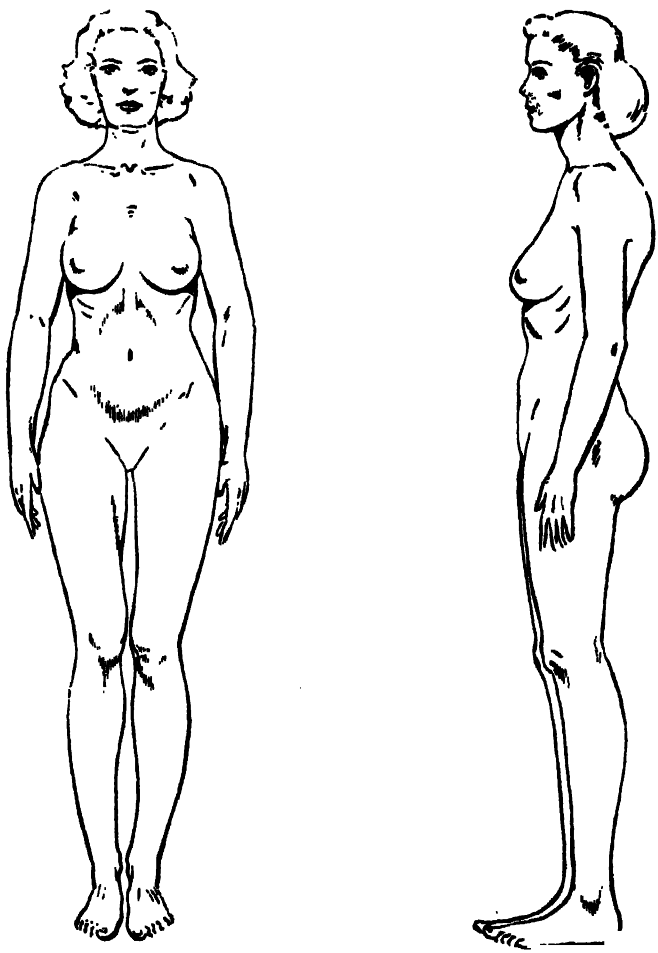


Fig. 13-6. Classic Female Somatotype 117.

subject at his greatest achieved distance from the ectomorphic pole or from the position of the 117 somatotype. The ponderal index is a measure of maximal achieved mass over surface. It can be reached at any age. Sheldon found the ponderal index helpful as an index of the third component, for as ectomorphy rises, nearly all diameter measurements fall sharply.

By referring to Figure 13-7 it is possible to calculate the ponderal index by placing a ruler on the appropriate height and weight of an individual and noting where it crosses the middle column ($H/\sqrt[3]{W}$). By noting this

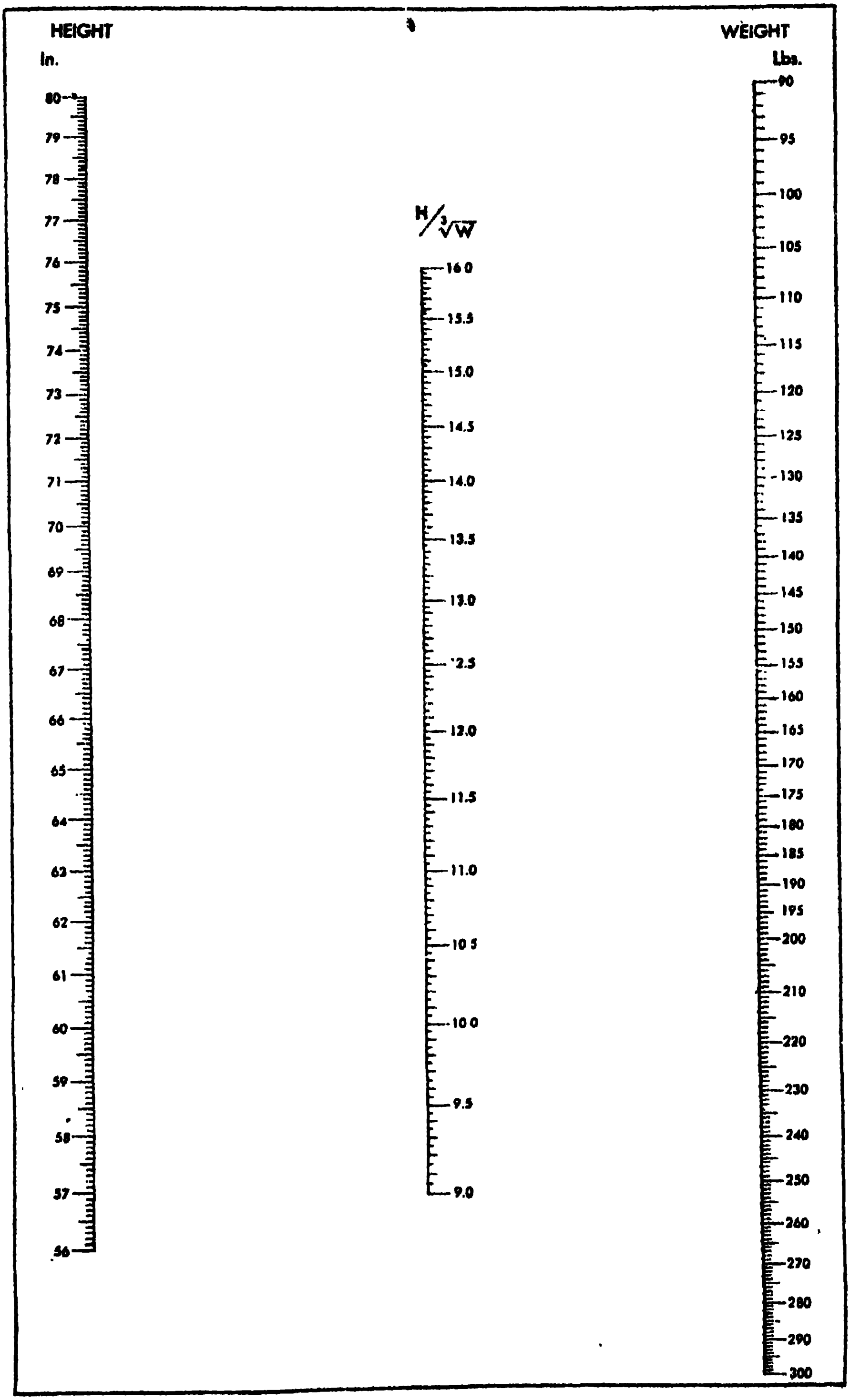


Fig. 13-7. Nomograph for Calculating Ponderal Index.

Ratio index	Height over cube root of weight						Age 18
	Somatotypes						
10.85	741						
10.90							
10.95							
11.00							
11.05							
11.10	731						
11.15							
11.20							
11.25							
11.30	721						
11.35							
11.40							
11.45	651 711	722					
11.50							
11.55							
11.60							
11.65		712					
11.70	561 641						
11.75							
11.80	471						
11.85							
11.90							
11.95		642					

Fig. 13-8. Chart for Calculating Somatotype from Ponderal Index, Age Eighteen. (From W. H. Sheldon, *Atlas of Men;* used by permission.)

Ratio index	*Height over cube root of weight*						Age 18
	Somatotypes						
12.00	631						
12.05	551						
12.10	371 461						
12.15							
12.20	621	632					
12.25	541	552					
12.30	271						
12.35		462 622					
12.40	361						
12.45		612					
12.50	171	272 542	623				
12.55	261 451						
12.60		362					
12.65		172	613				
12.70		452 532	543				
12.75							
12.80		262					
12.85		352 522	533				
12.90		162 442	453				
12.95			263				
13.00		252	353 523	534			
13.05			443				
13.10			163				

Fig. 13-8. (*Continued*)

Ratio index	Height over cube root of weight						Age 18
			Somatotypes				
13.15			433	354 524			
13.20			253 343	444			
13.25							
13.30				514			
13.35				254 434			
13.40				344			
13.45				424			
13.50				154	435 515		
13.55				244 334	345		
13.60					425		
13.65							
13.70					245 335		
13.75					325 415		
13.80					235		
13.85					145		
13.90						326	
13.95					225	236	
14.00							
14.05						316	
14.10						226	
14.15						136	
14.20							
14.25						216	

Fig. 13-8. (*Continued*)

Ratio index	Height over cube root of weight						Age 18
			Somatotypes				
14.30							
14.35						126	
14.40							
14.45							217
14.50							
14.55							
14.60							127
14 65							
14.70							
14.75							
14.80							117

Fig. 13-8. (*Continued*)

obtained index number in Figure 13-8, the probable somatotype is found. For example, a college boy 5 feet 8 inches tall and weighing 160 pounds would have a ponderal index of 12.55. In Figure 13-8 this index suggests possible somatotypes of 261 and 451. Because there is such a wide difference in the two possibilities, it is relatively easy to pick a type close to the visual estimate of the subject. The beginner, therefore, who has only a photograph or even a visual picture of the subject can begin right away to estimate the somatotype.

Another helpful piece of information to have on hand when estimating somatotype is the degree to which it is common in the population.[17] There are eighty-eight basic somatotypes, and the rarity for each has been determined by Sheldon. Figure 13-9 shows the approximate incidence per 1,000 persons.[18]

[17] We are dealing with the male population. Until additional female charts are completed, these figures can be used for women for estimation purposes, keeping in mind that the female constitution is more endomorphic throughout. For a distribution of 4,000 female somatotypes, see W. H. Sheldon. *Atlas of Men*, p. 13.

[18] Developed by W. H. Sheldon; used by permission. Actually, there are 713 somatotypes when all of the one-half combinations and overloaded components are included. When a somatotype is extra heavy (6 or 7) in two components, it is considered overloaded.

The *common somatotypes* run 31 to 60 per 1,000. Thus the 244, 343, 353, 443, 444, and 452 types are quite common. The *less common* run 11 to 30 per 1,000. The *scarce* run 4 to 10 per 1,000. The *very scarce* run 1 to 3 per 1,000. The *rare* occur less than 1 per 1,000. Thus, by consulting Figure 13-9 it will be found that somatotypes like 711, 117, 316, 515, 612, and 711 are going to be hard to find. In fact, in the college population they are practically nonexistent. Obviously, college performance is far too rigorous and demanding of such extreme constitutions, especially the mesopenes.

Somatotype	Incidence per 1,000	Somatotype	Incidence per 1,000	Somatotype	Incidence per 1,000
117	0.2	316	0.5	514	0.3
126	0.5	325	10	515	0.1
127	1	326	6	522	0.5
136	3	334	32	523	2
145	4	335	34	524	4
154	4	343	50	532	8
162	1	344	57	533	18
163	3	345	27	534	10
171	0.3	352	28	541	3
172	0.5	353	56	542	31
		354	37	543	17
		361	2	551	2
		362	18	552	6
		371	2	561	0.7
		372	3		
216	0.2				
217	0.4	415	0.5	612	0.2
225	3	424	4	613	0.5
226	12	425	5	621	0.5
227	2	433	9	622	3
235	20	434	22	623	2
236	15	435	14	631	2
244	40	442	25	632	7
245	24	443	60	641	2
252	4	444	42	642	5
253	35	451	4	651	0.5
254	28	452	48		
261	0.5	453	28		
262	10	461	2	711	0.1
263	14	462	6	712	0.3
271	1	471	3	721	0.7
272	3			722	2
				731	1
				732	3
				741	3

Fig. 13-9. The Approximate Incidence per 1,000 Persons for Each of the Eighty-eight Basic Somatotypes in the American Male Population. (Courtesy W. H. Sheldon.)

Anthropometric and Anthroposcopic Appraisals

For research purposes, where a very accurate somatotype figure is needed, it will generally be necessary to measure anthropometrically with the use of the complete photographic technique and IBM computer or substitute. The complete technique for doing this is described in detail in *Atlas of Men*.[19] Briefly, it consists in photographing the subject on three planes on a single film: front, side, and rear. A data board with the date and the subject's age, height, and weight is recorded in the picture. A good camera, proper lighting, positioning of subject, processing, and printing are very important. Experimentation in Oregon and Columbia medical centers has standardized this into routine procedure.

For the average teacher or student of constitutional psychology the above routine in its entirety is too time-consuming. For this person the *Atlas of Men* is especially useful. Here the obtained photograph can be compared through inspection with photographs of the eighty-eight somatotypes.

Simplified Method of Determining Somatotype

The easiest, and also the least accurate, method of somatotyping a person is simply to study the physical characteristics of the three components and then practice looking at persons in gym suits and bathing suits in order to select the primary and secondary components. Then do the following:

1. Select a subject to be somatotyped.
2. Decide on the primary component.
3. Rate it on the 1-to-7 scale.
4. Then estimate the secondary component on the 1-to-7 scale.
5. Do the same for the third component.
6. This will produce a three-digit somatotype, which may be named according to the descriptive classification already referred to.
7. Now figure the ponderal index (Figure 13-7) and relate it to possible somatotypes (Figure 13-9).
8. Finally, armed with one or two possible somatotypes, look through *Atlas of Men* to see how close you are. It is a full book of excellent pictures of male somatotypes with highly enlightening and sometimes amusing comments. Then make the final somatotype assessment.

The above eight steps work well. Even without *Atlas of Men* or the ponderal index, it is possible to estimate the somatotype close enough for teacher use with school children and college students. Without a great amount of training the teacher interested in somatotyping can learn to derive somatotype ratings through inspection. The author has subjected more than 1,200 college elementary education students to this technique.

[19] W. H. Sheldon, *op. cit.*, Appendix 3 (an elaborate handbook for somatotyping consisting of 1,175 examples of eighty-eight somatotypes with age-height-weight tables and curves for each).

They are fascinated by the topic and learn quickly how to estimate the components and apply the somatotype information directly with school children, keeping in mind that they do not have at their finger tips an absolute appraisal device. All this is not without foundation. Carefully trained workers have obtained reliability and objectivity coefficients of .92 using the inspection method.[20] Tanner also obtained high objectivity this way and found that trained observers agree to within half a rating on a 7-point scale in 90 per cent of cases.[21] Hawthorne, in checking objectivity, obtained coefficients ranging from .87 to .91.[22]

There is one other relatively simple anthroposcopic method which is more objective than the former technique. This is to use a rating scale for the components while observing the subject. Cureton, at the University of Illinois, was the first to employ subjective judgments, where the examiner rates the subject using a check list.[23] An adaptation of the routine is as follows:

1. Carefully study the morphological characteristics of each extreme body type.

2. Select a subject to be somatotyped and complete the following form according to the instructions that accompany it.

Subjective Somatotype Form

*Regional estimate by check list study**

Region of body	Endomorphy	Mesomorphy	Ectomorphy
I. Head, face, and neck	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
II. Thoracic trunk	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
III. Arms, hands, and shoulders	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
IV. Abdominal trunk	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
V. Legs and feet	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
Average of five regions	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7

Estimated somatotype (numerical) ....

Name of somatotype ....

* Rate each component separately on the 7-point scale by encircling the appropriate number.

a. Review the characteristics of each body type.

b. Consider one region of the body at a time and rate each component

[20] Thomas K. Cureton, "Body Build as a Framework of Reference for Interpreting Physical Fitness and Athletic Performance," *The Research Quarterly* (Supplement), 12:301–330, May, 1941.

[21] J. M. Tanner, "Reliability of Anthroposcopic Somatotyping," *American Journal of Physical Anthropology*, 12:257–260, June, 1954.

[22] Jesse J. Hawthorne, "Somatotyping and Its Relationship to Selected Motor Performance of College Men," unpublished doctoral dissertation, University of Texas, Austin, Tex., 1954.

[23] Thomas K. Cureton, *op. cit.*, p. 310.

(endomorphy, mesomorphy, ectomorphy) separately on the 7-point scale by encircling the appropriate number.

Example

(1) Rate pupil from 1-7 on Region I in endomorphy.
(2) Rate pupil from 1-7 on Region I in mesomorphy.
(3) Rate pupil from 1-7 on Region I in ectomorphy.
(4) Repeat for Regions II, III, IV, and V.

c. Average each of the encircled numbers for the five regions under endomorphy. This is the first component of the somatotype. Do the same for mesomorphy. This is the second component. Repeat for ectomorphy. This is the third component.

d. Place the three components in order. This is the estimated somatotype.

e. Name the somatotype.

Three Dimensional Approach

In an effort to make somatotyping more accurate for intensive research purposes, Sheldon and the personnel of his constitution laboratory have developed additional measures. A Distal Index, consisting of six distal measures, may be employed as a reciprocal of the Ponderal Index. In this way, both ponderosity and length are brought together mathematically to identify the somatotype.

In addition, and perhaps of greater value, is the Trunk Index. Using a planimeter to measure body surface area, somewhat as a geographer measures distances on a map, the research worker appraises the thoracic dorsal area (from the posterior crest of the ilium to the shoulder line across the neck) and the abdominal dorsal area (from the lower gluteal fold to the posterior crest of the ilium). This ratio, thoracic area/abdominal area, results in a 25-point scale in steps of 0.5 and is the Trunk Index.

An IBM chart is available for both sexes, showing height on the vertical axis and Trunk Index on the horizontal axis, from which the research worker may examine 716 possible somatotypes and pick the most likely one. Another set of useful charts are those in which Ponderal Index (vertical) is plotted against Trunk Index (horizontal), with the possible somatotypes falling in appropriate blocks according to the age of the subject. It is now possible, therefore, to differentiate between mesomorphy and endomorphy more accurately than ever before, for, instead of a two-dimensional approach to identification of body types, a three-dimensional approach is employed.

Other Indices Related to Somatotype

Space does not permit a discussion of other indices relating to Sheldon's body types; but there are several important secondary variables observable and measurable in the somatotype.

Dysplasia or d-index. Dysplasia is the degree to which there are different mixtures of the primary components in differing regions of the body. It is the sum of the differences in somatotype among the five body regions. The person with a short trunk and long legs would be an example of dysplasia. The d-index is greater in females.

Gynandromorphy or g-index. Gynandromorphy is the degree of bisexuality or presence of the physical characteristics of the opposite sex. The males that most resemble women are quite rare, whereas those low in g-index, highly masculine men, are common. The mesomorphic female with broad shoulders and narrow hips is frequently high in g-index, and the endomorphic males with broad, high hips also have a high g-index. The high g-index appears to be found in the male more than in the female. More men reach a 5 in endomorphy than women reach a 5 in mesomorphy.

Textural Component or t-index. Textural component is the degree of physical refinement or aesthetic quality of the subject.

Hirsutism or h-index. Hirsutism is the degree of luxuriance of body hair, a characteristic more common among the heavy, massive endomorphic mesomorphs.

Psychiatric Index. This is the degree to which a subject leans toward the three poles of psychotic reaction patterns: manic-depressive psychosis, paranoid schizophrenia, and hebrephrenic schizophrenia.

Temperament Index. This is the degree to which a person possesses the three primary components of temperament: viscerotonia, somatotonia, and cerebrotonia. It is sometimes referred to as the psychotype.

Scale for Temperament

In order to understand fully the body type approach to health behavior and motor activity, one should be familiar with Sheldon's temperament scale.[21] Here the structure is compared with behavior at a temperament level affording considerable insight into the interests and activities and aspirations of the subject according to the components.

In many respects the most far-reaching part of Sheldon's work with constitutional types concerns temperament. He has shown that the three types are temperamentally different. Just as there are extremes of constitution, there are extremes of temperament. These are referred to as *viscerotonia, somatotonia,* and *cerebrotonia*, respectively. Close and prolonged observation of many subjects, combined with an adaptation of the factor-analysis technique, resulted in the isolation of sixty descriptive or determinative traits, twenty for each of the first-order components of temperament. The scale for temperament, if carefully studied and related to body build, reveals considerable information (Figure 13-10).

The morphology of endomorphy is expressed through viscerotonic behavior; mesomorphy supports somatotonic behavior; and ectomorphy is as-

[21] William H. Sheldon and S. S. Stevens, *op. cit.*, p. 26.

sociated with cerebrotonic behavior In short, a specific structure is responsible for quantity and quality of function This is sometimes easy to appreciate After all, it is to be expected from daily observations that the roly-poly, pear-shaped endomorph is generally slow in movement and easily relaxed, that the athletic mesomorph is energetic and loves physical adventure, and

I *Viscerotonia*	II *Somatotonia*	III *Cerebrotonia*
1 Relaxation in posture and movement	1 Assertiveness of posture and movement	1 Restraint in posture and movement tightness
2 Love of physical comfort	2 Love of physical adventure	2 Physiological overresponse
3 Slow reaction	3 The energetic characteristic	3 Overly fast reactions
4 Love of eating	4 Need and enjoyment of exercise	4 Love of privacy
5 Socialization of eating	5 Love of dominating lust for power	5 Mental overintensity, hyperattentionality, apprehensiveness
6 Pleasure in digestion	6 Love of risk and chance	6 Secretiveness of feeling, emotional restraint
7 Love of polite ceremony	7 Bold directness of manner	7 Self-conscious motility of eyes and face
8 Sociophilia	8 Physical courage for combat	8 Sociophobia
9 Indiscriminate amiability	9 Competitive aggressiveness	9 Inhibited social address
10 Greed for affection	10 Psychological callousness	10 Resistance to habit, and poor routinizing
11 Orientation to people	11 Claustrophobia	11 Agoraphobia
12 Evenness of emotional flow	12 Ruthlessness freedom from squeamishness	12 Unpredictability of attitude
13 Tolerance	13 Unrestrained voice	13 Vocal restraint and general restraint of noise
14 Complacency	14 Spartan indifference to pain	14 Hypersensitivity to pain
15 Deep sleep	15 General noisiness	15 Poor sleep habits, chronic fatigue
16 The untempered characteristic	16 Overmaturity of appearance	16 Youthful intentness of manner and appearance
17 Smooth, easy communication of feeling, extraversion of viscerotonia	17 Horizontal mental cleavage, extraversion of somatotonia	17 Vertical mental cleavage, introversion
18 Relaxation and sociophilia under alcohol	18 Assertiveness and aggression under alcohol	18 Resistance to alcohol and other depressant drugs
19 Need for people when troubled	19 Need for action when troubled	19 Need for solitude when troubled
20 Orientation toward childhood and family relationships	20 Orientation toward goals and activities of youth	20 Orientation toward the later periods of life

Fig 13-10 The Scale for Temperament

that the frail, delicate ectomorph shows restraint both in posture and social address. A study of the scale indicates several instances of extreme temperamental differences of the constitutional components. In the endomorph, for example, eating is not merely a necessity as with the cerebrotonics; it is something to be thoroughly enjoyed. One is at home with persons in a way that the sensitive and often self-conscious ectomorph cannot possibly feel. In the somatotonic mesomorph, for example, the general tolerance of noise, the Spartan indifference to pain, and the overmaturity of appearance contrast with the ectomorph's general restraint of noise, hypersensitivity to pain, and youthful manner and appearance. Of course, persons combine all three components of temperament just as they do physical components. Thus the mesomorphic endomorph combines somatotonia and viscerotonia, with a greater strength in the latter, and is weak in cerebrotonia.

The morphological characteristics correlate quite well with the temperament scale. Sheldon's correlation between ectomorphy and cerebrotonia, for example, is .80. From the Grant study, an independent analysis conducted at Harvard University, it was found that there is a strong association between dominant ectomorphy and dominant cerebrotonia, with as high a proportion as 84 per cent of dominant ectomorphs being classified as dominant cerebrotonics.[25] This association was also statistically significant.

The components of structure and function are best viewed on a continuum. One overlaps the other from one end of the scale to the other. But along the continuum it is possible to see human traits of dependence, dominance, and detachment in the extremes. The viscerotonic traits of the endomorph seem to involve him in such a way that he is *dependent* on an environment of persons and things. He waits for a world on which he can depend. On the other hand, the traits of the cerebrotonic ectomorph are characterized by *detachment*. It is a movement away from excessive external stimulation, away from a demanding pushing world, and away from comfort and power. There is a movement toward the inner man and a greater awareness of oneself.

With the mesomorph, the tendency is to want to *dominate* the situation; there is often an unconscious desire and excitement in overcoming persons and things. In spite of environmental modifications in the traits of the somatotypes, there are still varying degrees of striving for dependence, dominance, and detachment. It shows up on the athletic field as well as in the classroom.

The Psychiatric Components

Sheldon devised a psychiatric classification that has a tremendous bearing on mental health.[26] It is too involved to study here. Briefly it relates extreme psychiatric behavior to component *weaknesses*. Thus, endopenia, mesopenia,

[25] C. Seltzer, F. L. Wells, and E. B. McTernan, "A Relationship between Sheldonian Somatotype and Psychotype," *Journal of Personality,* 16:431–437, June, 1948.
[26] William H. Sheldon et al., *The Varieties of Delinquent Youth,* New York, Harper & Brothers, 1949.

and ectopenia, when they appear in the somatotype, have a psychiatric implication. The perfect or complete manic-depressive is an ectopene (example, 551); the perfectly and completely paranoid schizophrenic is an endopene (example, 145), and the hebephrenic schizophrenic is a mesopene (example, 515). The latter would be almost entirely lacking in aggressive energy. This is what Jung might have called an archetype for the jettisoning personality, the person who can accept no responsibility, who in a mental institution might not even be able to feed himself. A clearer picture of this relationship is brought out as follows:

Body type	Psychiatric component I, manic-depressive	Psychiatric component II, paranoid	Psychiatric component III, hebephrenic
Endomorphy....	.54	−.04	−.25
Mesmorphy......	.41	.57	−.68
Ectomorphy.....	−.59	−.34	.64

Considering the variables that cloud the picture and are difficult to eliminate from comparison studies, the above coefficients of correlation are quite significant.

Sheldon has also suggested three types of delinquency based on morphology, all of which are related to endopenia, mesopenia, and ectopenia. These will be of special interest to recreation workers and personnel in adapted physical education. They are covered through a study of individual cases in *The Varieties of Delinquent Youth.*

APPLICATION OF SOMATOTYPE TO HEALTH AND PHYSICAL EDUCATION

Few educational areas stand to profit as much from knowledge of constitutional psychology as health and physical education. With body build related to physical and mental health, motor fitness, motor ability, and success in athletics, it is imperative that workers in health guidance, athletic coaching, and other related fields be versed in the many implications and practical applications inherent in the study.

In *Man the Unknown,* the 1912 Nobel Prize winner, Alexis Carrel, relates body morphology to physical fitness[27]:

> Resistance to disease, work, and worries, capacity of effort, and nervous equilibrium are the signs of the superiority of man.... Each man is characterized by his figure, his way of carrying himself, the aspect of his face. Our outward form expresses qualities, the power of our body, and of our mind.... The state of the organs is revealed by the aspect of the body.... In fact, people who belong to different morphological classes—for instance, to the cerebral, digestive, muscular, or

[27] Alexis Carrel, *Man the Unknown,* New York, Harper & Brothers, 1935, p. 62.

respiratory types—are not liable to the same organic or mental diseases.... Each man bears on his face the description of his body and soul.

In medicine there are three well-defined and essentially interrelated variables: the human subject, the disease process, and the special environmental stress. There is endless work in each of these variables. The behaviorists and environmentalists have gone on so long unchecked that, if one is not careful, one will be led to believe that heredity and constitutional make-up are simply incidental variables when it comes to the actions of boys and girls in school. This obviously is not true, for human behavior at any level *tends* to be related to constitutional make-up. Nowhere is this better brought out than with physical and mental diseases, in both clinical and nonclinical manifestations.

Health Implications of the Somatotype

The three extreme somatotypes respond differently when exposed to similar variables, such as pathogenic organisms, diet, fatigue, and rest. Possibly the health status of the extreme body types is as different as the behavior of the hippopotamus and the greyhound.

Consider the endomorph. He is usually a well-adjusted, relaxed, and amiable person. There is an evenness of emotional flow and tolerance. This is not generally the type of school child that is likely to fly into a tantrum of anger. Moreover, because neither he nor the extreme ectomorph particularly relish physical work, they are not prone to the accidents experienced by the mesomorphic types. For example, it is not the endomorph or the ectomorph that takes off on a mountain-climbing expedition.

Although the endomorph has a superb digestive apparatus and rarely suffers from indigestion, he is the one that has gall bladder and cardiovascular difficulties. His weight is a handicapping factor in many respects. Because of it, his longevity is considerably reduced. The adverse effect of overweight in each sex makes itself felt at every period of life. Mortality from cardiovascular-renal diseases among endomorphic persons is about 50 per cent above standard experiences.[28]

Among endomorphs, especially those rated 5, 6, or 7, liver cirrhosis is high, and diabetes mellitus is a most striking penalty disease. These persons have hypertension, degenerative arthritis, and abdominal hernia. They are poor surgical risks. Extreme endomorphic females have more complications of pregnancy, with operative procedures twice as common at delivery; and the stillbirth rate is 1½ to 2 times as great among the massive endomorphs, endomorphic mesomorphs, and mesomorphic endomorphs. Toxemia of pregnancy is frequent in short-legged, heavy, endomorphic women.[29]

[28] Metropolitan Life Insurance Company, "Overweight Shortens Life," *Statistical Bulletin*, vol. 32, no. 10, October, 1951; see also "New Weight Standards for Men and Women," *Statistical Bulletin*, vol. 40, November-December, 1959.

[29] George Draper, Clarence W. Dupertius, and J. L. Gaugher, Jr., *Human Constitution in Clinical Medicine*, New York, Paul B. Hoeber, Inc., 1944, p. 22.

It should be pointed out that heaviness or massiveness of body build frequents both endomorphs and mesomorphs. Very often researchers fail to differentiate between types that are essentially obese, hefty endomorphs and huge mesomorphs who have in later years added considerable weight.[30] In short, an overweight mesomorph is not the same constitution at all as an endomorph of the same weight. In this respect, Sheldon has established height-weight tables for each of the known somatotypes based in part on weight changes with age.[31] As might be expected, over the years the perfectly normal endomorph gains considerable weight, whereas the mesomorph gains in general, depending on the degree of endomorphy as a secondary component. The ectomorph gains the least. The more extreme the ectomorphic component, the less the gain. This is nicely illustrated in Figure 13-11, where three ectomorphic somatotypes are structured for weight changes over the years from eighteen to sixty-three.[32]

By examining the three height-weight tables, 261, 262, and 263, it will be noted that ectomorphy has increased from 1 to 3, respectively, whereas endomorphy and mesomorphy remain the same. Note further that at age eighteen, with a height of 61 inches, the weight decreases for each somatotype: 114, 108, and 104, respectively. Further analysis indicates that the 261 increased in weight from 114 to 135 pounds by age sixty-three, a total of 21 pounds. The 262 increased only 19 pounds and the 263 only 13 pounds. This suggests that the weight of school children and adults may have more to do with individual morphology and its refined classifications than was previously considered.

The mesomorph is the vigorous, extroverted type, the physically courageous person who possesses the muscular strength, endurance, and agility to enjoy physical adventure and possibly suffer the greatest number of accidents. This is the unrelenting, driving temperament who frequently puts forth tremendous physical efforts to reach a goal. In his efforts to move ahead, he possesses an almost Spartan indifference to pain, especially when compared with the overly sensitive ectomorph. Here is the "peptic ulcer type" and "duodenal ulcer type," especially when he leans toward ectomorphy. He also has more appendicitis and, when in combination with endomorphy, more poliomyelitis.

Polio is quite rare among extreme ectomorphs. Mesomorphy is somewhat associated with longevity, in that mesopenes are the shortest-lived somatotypes.

The ectomorph, in many respects, is at the other end of the pole from endomorphy when health behavior is considered. Medically, the extreme ectomorphic boys and girls, even at a primary level age, seem to have more

[30] Carl E. Willgoose, "Health Characteristics of the Somatotypes," *The Physical Educator*, 12:100–102, October, 1955.

[31] See eighty-eight height-weight tables in *Atlas of Men*.

[32] William H. Sheldon, personal correspondence.

Age*	*Somatotype 261:* height (in inches)														
	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
18	114	120	126	132	138	145	152	159	166	173	180	188	195	203	212
23	119	124	131	137	144	151	158	165	172	180	187	193	204	212	221
28	123	129	135	142	149	156	163	170	178	186	193	202	210	219	228
33	126	133	139	146	153	160	168	175	183	191	199	208	216	225	234
38	130	136	143	150	157	165	172	180	188	196	205	214	222	232	241
43	132	139	146	153	161	168	176	184	192	201	209	219	227	237	247
48	134	141	148	155	163	170	178	187	195	204	212	222	230	239	249
53	135	142	149	156	163	171	179	187	195	204	212	222	230	239	249
58	135	142	149	156	164	171	179	188	196	205	213	222	230	240	250
63	135	142	149	156	163	171	179	187	196	205	212	221	230	239	249

Age	*Somatotype 262:* height (in inches)														
	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
18	108	114	119	125	131	137	144	151	157	165	172	179	186	194	202
23	114	120	126	132	138	145	152	159	166	174	182	189	197	205	213
28	117	123	129	135	142	149	156	163	170	179	186	194	202	210	220
33	120	126	132	138	145	152	159	167	174	183	190	198	206	215	224
38	123	130	136	142	149	156	164	171	179	188	196	204	212	222	231
43	125	131	137	144	151	158	166	173	181	190	198	206	214	223	232
48	127	133	140	146	154	161	168	176	184	193	201	209	218	227	237
53	127	134	140	147	154	161	169	177	185	194	202	210	218	228	237
58	127	134	140	147	154	161	169	177	185	194	202	210	218	228	237
63	127	133	139	146	154	161	168	176	184	193	201	209	218	227	236

Age	*Somatotype 263:* height (in inches)														
	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
18	104	109	115	120	126	132	138	145	151	159	165	172	179	186	194
23	108	113	118	124	130	136	143	149	156	164	170	177	184	192	201
28	110	116	121	127	133	140	146	153	160	168	175	182	189	197	206
33	113	118	124	130	136	143	149	157	164	171	179	186	193	201	210
38	114	120	126	132	138	145	152	159	166	174	181	189	196	204	213
43	116	121	127	133	140	147	153	161	168	176	183	191	198	207	216
48	117	123	129	135	142	149	156	163	170	179	186	193	201	210	219
53	117	123	129	135	142	149	156	163	170	179	186	193	201	210	219
58	117	123	129	135	142	149	156	163	170	179	186	193	201	210	219
63	117	123	128	135	142	149	156	162	170	179	185	193	201	209	219

* 18 represents average of 16–20 span; 23 represents average of 21–25, etc.

Fig. 13-11. Examples of Weight for Age, Height, and Somatotype.

troubles of the upper respiratory tract and skin. They are so sensitive to all stimuli that they are easily nauseated, have dry scalps, violent head colds, and sore throats. No wonder Hippocrates called them the respiratory type. Their sensitivity, self-consciousness, and hyperattentiveness is reflected throughout the total nervous system. Thus, they are easily bothered by insect bites, on the one hand, and noise, disorganization, and confusion, on the other. Poor sleep habits and chronic fatigue is the all too common result. Moreover, they are seen with numerous postural defects and readily swollen mucous membranes from acute streptococcal infections. They are simply a different breed of cat than the extreme endomorphs in society. Although infectious diseases bother them, they do not have as much tuberculosis as was formerly believed. Tuberculosis seems to frequent all types. They are the longest lived. This is all the more true if they are short in stature, a fact noticed by William James years ago and commented upon by Sheldon. This longevity may be because of or despite their great sensitivity, a characteristic that may penalize them in childbearing, for the miscarriage rate for extreme ectomorphs appears to be three times as great as that for the other types.[33]

Combinations of components are the most revealing when it comes to health implications. Endomorphy and mesomorphy, for example, produce cancer, a disease extremely rare in ectomorphs. Women who have cancer of the breast and uterus are nearly all solidly built, hefty women who are quite ectopenic (low ectomorphy).[34] The mesomorphic ectomorph and ectomorphic mesomorph experience ulcers, and endomorphic mesomorphs and mesomorphic endomorphs have coronary disease. In this latter disease the pathological process relates directly to the constitution. Clinical experience suggests that the coronary arteries in some persons are as acutely responsive to strong emotional stimuli as they are to other strains.[35] Prolonged fear, apprehension, worry, anxiety, hatred, or anger are more significant here than in other builds. These persons might express themselves as John Hunter did: "My life is in the hands of any rascal who chooses to annoy and tease me."

Gertler and White showed in their research that the extremes of the three components hardly ever had coronary heart disease. It was learned that "coronary heart disease is significantly associated with a specific body habitus rather than with body weight."[36] It was also clearly demonstrated that the endomorphic mesomorph, a heavy, muscular person, is most prone to the disease, whereas the lean ectomorph is the least prone to it.

Finally, a word on mental health implications. More and more evidence

[33] From obstetrical research studies, Central New York Project, Frank C. Meyer, M.D., and Carl E. Willgoose.

[34] Brought out by George Draper, Clarence W. Dupertius, and J. L. Gaugher, Jr., *op. cit.*, p. 24, and William H. Sheldon et al., *op. cit.*, p. 788.

[35] George Draper, Clarence W. Dupertius, and J. L. Gaugher, Jr., *op. cit.*, p. 37.

[36] Menard M. Gertler and Paul Dudley White, *Coronary Heart Disease in Young Adults*, Cambridge, Mass., Harvard University Press, 1954.

suggests that, for a student's actions and attitudes to be meaningful to the researcher or teacher, they probably should be interpreted against the frame of reference of a well-described physical constitution. There is no separating the physical and the mental. We find only structure and behavior, which seem to make a functional continuum.

The health educator must study structure and psyche together. The psychoneuroses of poor mental health and the psychoses of the seriously deranged relate to biological structure. The three components of body build represent three poles around which to group this mental behavior.[37] The manic-depressive, who is emotionally up or down with mania or melancholy,

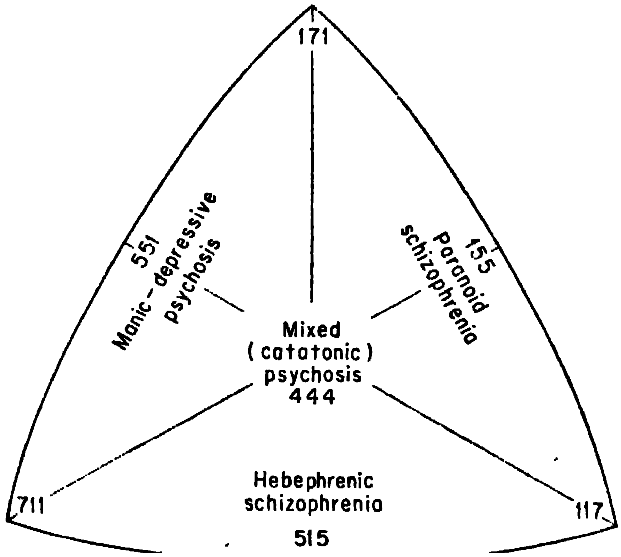


Fig. 13-12. Body Type and Psychotic Extremes.

possesses an entirely different body build than either the paranoid or hebephrenic schizophrenic. The manic-depressive is found among the endomorph-mesomorph combinations, who are very low in ectomorphy (ectopenes, 551, 451, 541, etc., see Figure 13-12).

The combination of strong mesomorphy and ectomorphy with very low endomorphy (endopene) generally results in paranoid schizophrenia. The 155, 145, or 154 is aggressive with somatotonic drive and delusions of persecution. He possesses the physical constitution to put forth great (wasted) energy in fighting against something. The hebephrenic schizophrenic, on the other hand, is extremely weak in constitution, having very little mesmorphy (mesopene). This 515, 425, or 415 type gives up easily, has a youthful mind, loses interest, and regresses to an infantile level.

As psychiatric cases, the midrange somatotypes, 343, 443, 344, etc., have

[37] William H. Sheldon et al., *op. cit.*

fallen in the mixed psychosis or catatonic category, difficult to categorize compared with the extreme body types.

In order for this psychiatric research to be valuable in education and mental health, the user must realize that certain constitutional types may *lean* toward slightly deranged or neurotic behavior. *Many never become a "statistic" because environmental influences are strong enough to overcome any tendencies toward psychotic behavior.* Thus, the school child or young adult who shows signs of maladjustment needs help. The self-critical, over-sensitive cerebrotonic needs encouragement in the sports where he feels the most comfortable: individual sports. Putting him in group activities to cultivate group consciousness might drive him out of sports altogether, losing the physical education teacher an opportunity to be of service. There is little doubt that what is one man's meat is another man's poison.

In view of what is known about the general health characteristics of the somatotypes, there appear to be several important implications for health education and physical education personnel in schools and colleges.

1. Health records should reflect the somatotype rating.
2. Studies of growth, height and weight, and nutrition of school children should consider the strength of somatotype components.
3. The punishment of children, whether light or severe, verbal or physical, needs to be related to the constitutional sensitivity of the child.
4. The school environment, both physical and personal, may mean more to the success of one type than another. This will involve the health co-ordinator and his problem children of subpar physical and mental health.
5. School health examinations and screening devices will need interpreting in the light of body build.

Physical Fitness

It is somewhat of an axiom that those who *can*, do: by extension, those who cannot do are unable. Thus the capacity for physical activity underlies both the quality and the quantity of action. It is the mesomorphs that possess the structure for vigorous and prolonged efforts of strength, power, and endurance. This is also the type that matures early and is ready to engage in sports and athletics ahead of contemporaries.

Early in his research Sheldon found low negative correlations between peripheral strength and endomorphy (—.33) and ectomorphy (—.19) and a positive association with mesomorphy (.64).[38] About the same time Seltzer and Brouha, of the Harvard University Department of Hygiene, demonstrated that the mesomorphs had vastly better physical fitness ratings on the Harvard Step Test after twelve weeks of training than did the endomorphs.[39] Earlier Cureton and Hunsicker found that, in general, mesomorphs and

[38] William H. Sheldon and S. S. Stevens, *op. cit.*, p. 405.

[39] Carl C. Seltzer and Lucien Brouha, "The Masculine Component and Physical Fitness," *American Journal of Physical Anthropology*, 1:21–28, March, 1943.

ectomorphic mesomorphs score higher in physical fitness tests than do other somatotypes.[40] It was recommended that physical fitness be interpreted in terms of body build. This same point was made later by Willgoose and Rogers in a study where it was found that the highest PFI score was attained by the mesomorph-ectomorph group and the next highest mean score by the mesomorph group.[41] Although the study was based on a limited number of cases the relationship between the components is useful (see Figure 13-13). Here as mesomorphy increases, the PFI score goes up. Where the predominantly mesomorphic person had an endomorphic component of less than 4, the average PFI score was 106.9; but when the endomorphic component was 4 or more, the mean fitness score dropped to 80.5. Where endomorphy is definitely a limiting factor, ectomorphy may mean very little by itself. Fitness of the ectomorphic student seems to depend on the amount of mesomorphic component displayed.

Jones, in a study of eighty boys, found a correlation of .61 between strength and mesomorphy when height and weight had been eliminated.[42] Sills also categorized endomorphy as a limiting factor and mesomorphy as a positive factor in tests of strength and motor ability.[43] Later this work was expanded to demonstrate again that excess weight is a handicap to endomorphs and that insufficient strength is a handicap to ectomorphs.[44] On the basis of such findings, Sills and Mitchem constructed four T-score tables for thirteen categories of somatotypes for performances on a four-item physical fitness test battery used at the State University of Iowa.[45]

In applying this information to somatotype and physical fitness, one is faced with two choices. Either new norm tables must be constructed for the numerous fitness tests, or existing norm tables must be interpreted in the light of individual body build. Bookwalter, after studying the test scores on the Indiana Physical Fitness Test of elementary school boys who varied in body build according to the Wetzel Grid, advised the establishment of norms so comparable sizes and shapes can be related to physical fitness.[46] True, it hardly seems fair to compare the endomorphic and ectomorphic extremes to

[40] Thomas K. Cureton and Paul Hunsicker, "Body Build as a Framework of Reference for Interpreting Physical Fitness and Athletic Performance," *The Research Quarterly* (Supplement), 12:300 330, May, 1941.

[41] Carl E. Willgoose and Millard L. Rogers, "Relationship of Somatotype to Physical Fitness," *Journal of Educational Research*, 42:704-712, May, 1949.

[42] Harold E. Jones, *Motor Performance and Growth*. Berkeley, Calif., University of California Press, 1949, p. 181.

[43] Frank D. Sills, "A Factor Analysis of Somatotypes and Their Relationship to Achievement in Motor Skills," *The Research Quarterly*, 21:424-437, December, 1950.

[44] Frank D. Sills and Peter W. Everett, "The Relationship of Extreme Somatotypes to Performance in Motor and Strength Tests," *The Research Quarterly*, 24:223–228, May, 1953.

[45] Frank D. Sills and John Mitchem, "Prediction of Performances on Physical Fitness Tests by Means of Somatotype Ratings," *The Research Quarterly*, 28:64–71, March, 1957.

[46] Karl W. Bookwalter, "The Relationship of Body Size and Shape to Physical Performance," *The Research Quarterly*, 23:271–279, October, 1952.

a norm table established on the total range of physiques. Yet it would be a time-consuming job to amend all norm tables. It appears to be much more practical to *interpret* fitness scores in the light of the somatotype. It is not difficult to relate the estimated somatotype to the performance. Here is where health, guidance and adapted physical education personnel can profit. Below-average performance of the endomorphic and ectomorphic types is to be expected, and the mesomorphic variety should do very well. PFI scores, for example, above 85 or 90 for the 5, 6, and 7 endomorphic type are quite

Descriptive components	*Mean PFI score*
Extreme endomorph (5^2, 6, 6^2, 7)	68.4
Mesomorphic endomorph	71.2
Endomorph-mesomorph (443, 552, 4^24^22)	77.7
Endomorphic mesomorph	
Endomorphic component 4 or more	80.5
Endomorphic component less than 4	106.9
Extreme mesomorph (5^2, 6, 6^2, 7)	112.8
Ectomorphic mesomorph	103.8
Extreme ectomorph (5^2, 6, 6^2, 7)	102.
Mesomorphic ectomorph	105.6
Mesomorph-ectomorph (144, 155, 244)	123.6
Equal components ($3^23^23^2$, 444)	100.3
Mesopene (434, 426, 315)	72.3

Fig. 13-13. Breakdown of Somatotype into Descriptive Components with Corresponding Mean PFI Score.

satisfactory. But such a score for the comparably weighted mesomorph would be quite substrength and suggest a thorough and more revealing physical examination of the subject.

Motor Ability and Athletic Performance

Motor ability follows much the same pattern as physical fitness. In the combination of components the mesomorphic element is significant. This shows up in physical education classes. It has been demonstrated to be present among student physical education majors.[47]

There is extensive literature on the relationship of body build to athletic performance. Cureton has outlined much of this.[48] Tyrance examined flexibility and showed that, when significant flexibility and body type variables are compared, the association indicates that as the mesomorphic com-

[47] John Compton Thompson, "An Analysis of the Factors Affecting the Achievement of Undergraduate Men Majoring in Physical Education at the State University of Iowa," *The Research Quarterly,* 23:417–427, December, 1952.

[48] Thomas K. Cureton and Paul Hunsicker, *op. cit.,* p. 2; see also Thomas K. Cureton, "Relation of Physical Fitness to Athletic Performance and Sports," *Journal of the American Medical Association,* 162:1139–1149, Nov. 17, 1956.

ponent increases, neck flexibility also increases.[49] Pollock demonstrated that the relationship between a preponderance of mesomorphic tissue and athletic ability, motor ability, and strength was of a higher order than the relationships of endomorphic or ectomorphic tissues to the physical traits.[50] Furthermore the mesomorphic ectomorphs proved to be better than the mesomorphic endomorphs.

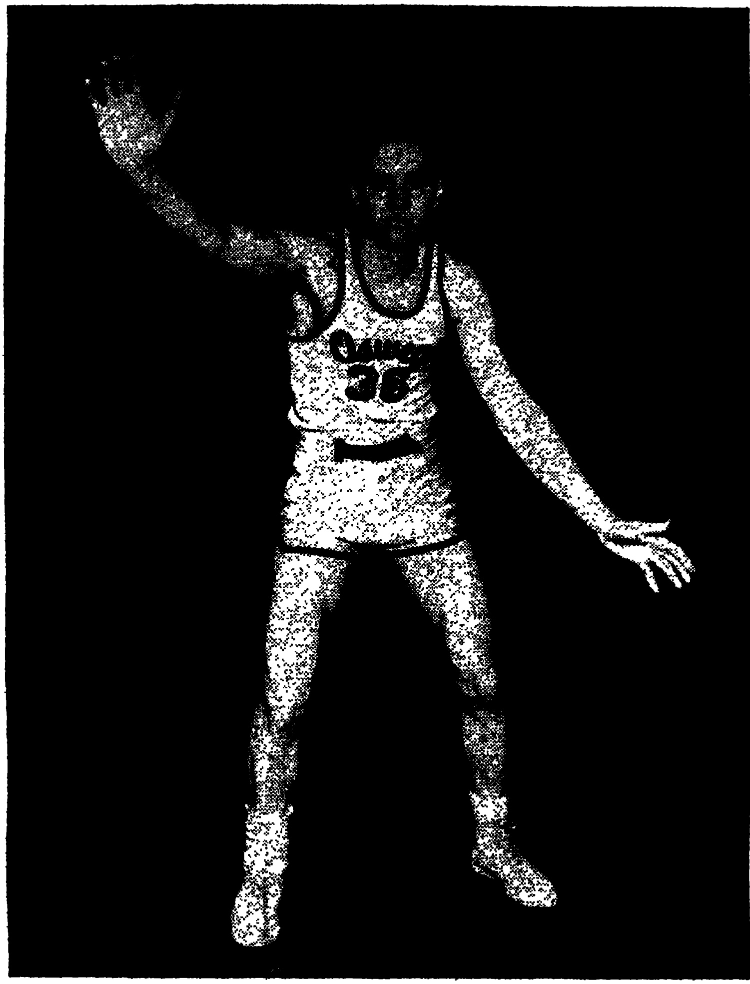


Fig. 13-14. High Mesomorphy of the Ectomorphic Components, S-25²3.

Studying pictures of athletes indicates numerous implications for physical education and athletic coaching (see Figures 13-14 and 13-15).[51]

[49] Herman J. Tyrance, "Relationships of Extreme Body Types to Ranges of Flexibility," *The Research Quarterly,* 29:349–359, October, 1958.

[50] Daniel J. Pollock, "Relationships between Somatotypes and Certain Measures of Personality, Interests, and Physical Activities of College Men," unpublished doctoral dissertation, University of Texas, Austin, Tex., 1956.

[51] Carl E. Willgoose, "Body Types and Physical Fitness," *Journal of Health, Physical Education and Recreation,* 27:26–28, September, 1956.

In Hawthorne's study of college men, the mesomorphic trait correlated above .67 with leg strength, and the ectomorphic trait correlated well with motor ability.[52] Both types excelled on tests of agility and reaction time. These were not extreme types; that is, these were the performances of mesomorphic ectomorphs and ectomorphic mesomorphs. Mesomorphic ectomorphs correlated .43 with the Brace test and .51 with the vertical jump.

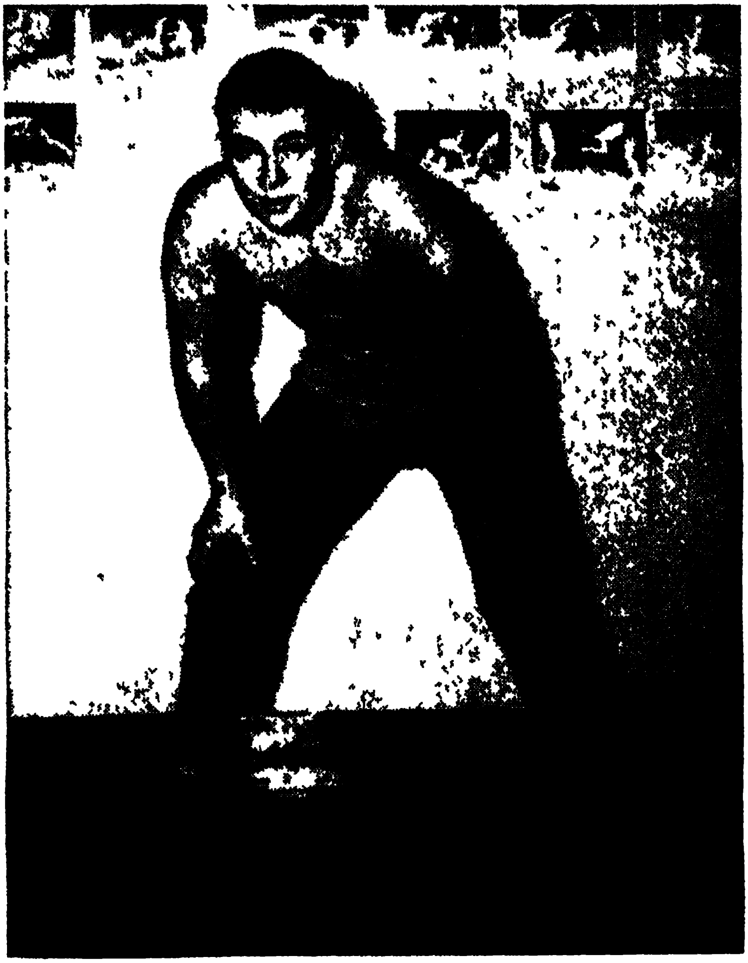


Fig 13-15 High Mesomorphy of the Endomorphic Component, S-45¹1².

The endomorphic trait correlated —.84 with the Brace test The findings of Chalkley, at the University of Maryland, are similar.[53] The athletic group tended to follow a pattern of body build. Track athletes and basketball players leaned more toward ectomorphy than did the swimmers and

[52] Jesse J. Hawthorne, *op. cit*, p. 66.
[53] Raymond E. Chalkley, "A Study of the Sports Activity of Young Adult Males and Its Relationship to Somatotype and Selected Anthropometrical Measurements," unpublished master's thesis, University of Maryland, College Park, Md., 1958.

wrestlers, who possessed a fairly high degree of mesomorphy In this study the extreme somatotype ratings were made subjectively by three judges Experience suggests that the combination of components in the somatotype is more revealing than the extremes An ectomorph with a 5 or 6 rating means little in terms of performance until the secondary component is known The 146 and 235 are quite different Yet if only the third component was checked, their differences might disappear by averaging out

Physical Education Implications

The three extreme somatotypes behave differently when exposed to the same educational stimuli In physical education classes the endomorph is often called "fatty" He goes through the motion of physical education activities but doesn't really enjoy them He is slow and awkward, lacks agility and strength and enters the activity chiefly to be part of the group

From the physical education point of view, the extreme endomorphs and ectomorphs present one problem after another It might well be a serious mistake to try to make mesomorphic gymnasts out of them Neither type feels comfortable in the gymnasium Whereas the endomorph enters group games easily, it is the ectomorph that prefers to remain detached and watch

The teacher who understands body types will attempt to control group situations by holding down the exuberance of the mesomorphs and encouraging the shy and sensitive persons to enter the group activity and become part of it It will always be a bit difficult to get the ectomorphic child to enter wholeheartedly into the group activity, for he tends to prefer a small circle of friends and to participate in quiet activities which do not involve large numbers

The potential fitness of the endomorph is limited especially when he is compared to a norm table that has been built on the performances of extreme mesomorphs and flyweight ectomorphs Yet this viscerotonic endomorph has little difficulty cooperating in a sporting situation

He is good to have around, for he often sees the humor in his lack of physical ability He will learn, therefore, some recreational skill, such as fly casting, shuffleboard, table tennis, or horseshoes With a higher degree of mesomorphy, he will begin to lean toward team sports and enjoy participating in them

The extreme ectomorph, on the other hand, may steer clear of gymnasiums unless his mesomorphy component is above a 3 or 4 When he exercises, he often prefers nonteam situations He goes out for track and field events, fencing, tennis and other single or dual activities

Extreme mesomorphs make good physical education students and top athletes One can examine any number of persons ranking high or low in the two components of endomorphy and ectomorphy and tell a great deal about their potential motor and athletic ability by the degree of mesomorphy present in the structure The mesomorphs need, and very much love, pure

exercise. Whereas the slim ectomorph, standing in a gym suit, is extremely sensitive to his surroundings and feels stark naked, the mesomorph tends to throw out his chest and say, in effect, "just look at me" or "watch me move about."

Implication for Athletic Coaching

Men and women who spend hours training youths to perform athletic skills will find that the delicate relationship between the three somatotype components is a valuable area to become familiar with. Sheldon points out that a football coach who cannot distinguish between a 5^2 and a 6 in mesomorphy may have trouble winning games.

At least half the coach's success in athletics is picking the right participants to train in the beginning. Although poorly constituted men become fair players in a sport, the champion athletes are rich in the constitutional factors supporting above-average motor ability. Merely exposing the endomorph to a vast amount of training will not measurably improve agility and reaction time.

In many respects, athletes are born and not made. From the view of kinesiology and body mechanics, the V structure of the mesomorph (broad shoulders and narrow hips) is "engineered" for high functional motor ability. The pear-shaped structure of the extreme endomorph or the beanpole physique of the ectomorph presents one problem after another when motor ability is involved.

In the end, it seems to be the combinations of these constitutional components that affords some insight into human behavior on the athletic field or in the gymnasium. Certainly, distance swimmers differ decidedly from wrestlers, and basketball centers differ considerably from football quarterbacks.

It will be helpful to the coach to note that the somatotypes differ in their ability to train down to an ideal playing weight. Generally, those men who gain weight fast when less active lose more when in training. The endomorphic mesomorphs, who make good athletes, will train down in weight about 12 per cent when in shape.[54] On the other hand, should the ectomorphic mesomorph lose this much weight in training, he might significantly reduce his power and endurance.

Within a certain group of somatotypes, there are similar varieties of athletes, which appear as in Figure 13-16.[55]

Of course, there is some overlapping in the above structure, but as a rule, the hippopotamuslike endomorphic combinations and the greyhoundlike ectomorphic combinations have quite different interests and abilities in

[54] For specific somatotype height-weight tables for men up to sixty years of age, see Sheldon's *Atlas of Men*.

[55] Neither the extreme endomorphs nor the extreme ectomorphs consider motor activity seriously enough to engage in athletics. They are, therefore, omitted here.

athletics. Physically, it is easier for the extreme ectomorph to take up skiing than it is for either the extreme endomorph or mesomorphic endomorph. As a rule, however, the extreme ectomorph will have little interest in anything so vigorous as skiing. What probably appeals to him about the sport is the aesthetic beauty of winter and an activity in which he can remain somewhat detached from the group.

By the same token, the mesomorphic endomorph, dependent on other persons for security, may take up skiing under great physical odds in order

Mesomorphic endomorphs (S-Types: 631, 532, 541, 542, 543)	Endomorphic mesomorphs (S-Types: 452, 361, 462, 451, 453)	Extreme mesomorphs (S-Types: 171, 162, 262, 172, 252)	Ectomorphic mesomorphs (S-Types: 253, 254, 163, 164, 265)	Mesomorphic ectomorphs (S-Types: 235, 126, 136, 145, 146)
Table tennis Floating (swimming) Croquet Fly and bait casting Bowling	Baseball Football (lineman) Heavyweight boxing Heavyweight wrestling Swimming Soccer (backs) Ice hockey (backs) Weight tossing	Sprints Basketball Middleweight boxing Middleweight wrestling Quarterbacks Football (backs) Divers Tumbling Lacrosse Soccer (forwards) Ice hockey (forwards) Handball	Lightweight wrestling Long-distance running Tennis Gymnastics Weight lifting Javelin Pole vault High jump Fencing Badminton Skiing Jockey	Bicycling Cross country Table tennis Basketball center (short periods) Archery Also many athletic games, except those requiring weight and sheer strength

SOURCE: Carl E. Willgoose, "Body Type and Physical Fitness," *Journal of Health, Physical Education and Recreation*, 27:26–28, September, 1956.

Fig. 13-16. Body Types and Athletic Performance.

to feel that he is a part of an outing club or skiing group. His success may be very satisfying to both the group and himself.

To analyze the somatotypes further, it is enlightening to observe the motor actions of various combinations of components. Although the mesomorphic component is the driving force in the structure, it is the strength of the secondary component (endomorphy or ectomorphy) that reveals much about physical abilities. Extreme ectomorphs do little physical work, but when some degree of mesomorphy is present in the second component, the ectomorph probably will enjoy some physical movements.

For example, the 126 might well engage in hiking over a nonmountainous terrain. The 136, however, might venture forth into cross-country competi-

tion. The 145 or 146 would hike over mountainous territory, for he (like the 154) is swift, lean, lithe, somewhat tight and strained, and has ability in noncontact activities such as tennis and long-distance running. The 236 somatotype seems to possess excellent coordination, is often tall, and is good at distance running and playing center on a basketball team. He lacks the stamina to bounce about in a basketball game, so he plays center for short periods at a time. Sheldon says that if cherry picking were a competitive sport, a 236 might win a national championship.

The 5s and 6s in mesomorphy are interesting to compare. The 162 is an efficient fighting machine that can train down pretty well. He, like most of the ones in endomorphy, is light. With strength and agility, a short 162 weighing 110 pounds might train down to make a good jockey or a lightweight wrestler. To be a first-rate boxer, he would need a little more endomorphy.

In one of Sheldon's early works, he pointed out that although the 171 or 172 is all-powerful, rugged, and fast and might look like Dick Tracy, Lil' Abner, or Superman, he usually lacks the endomorphic padding that makes a durable fighter, one who can stand up in contact sports. Also this somatotype is not very "loose." He is, perhaps, too low in endomorphy to relax easily.

The college and professional baseball, football, boxing, and heavyweight wrestling men seem to be "rendered elastic" by a 2 or 3 or even 4 in the first component. Endurance itself seems to be based on the ability of the athlete to relax between sustained movements. Flexibility and elasticity are related to relaxation and are demonstrated in pitchers and fielders alike who have some fair degree of the first component. Observing the two or three dozen professional baseball players moving about the playing field, one will find almost without exception a number of 2s, 3s, and 4s in endomorphy. Although persons in this group add weight and encounter cardiovascular difficulties as they grow older, they are hard to beat in their prime.

It can readily be seen from the above discussion that the weight of the secondary component is especially revealing when considering athletic performance. The great powerful athletes, comparable to the mesomorphic Zeus and Hercules, usually have endomorphic and ectomorphic support. For example, the 354 is a healthy, long-lived fellow seldom seen in clinics. He is often tall, is an excellent athlete, and has enough endomorphy to keep relaxed, which nicely counterbalances his rather high degree of ectomorphy and cerebrotonic sensitiveness.

Sheldon reports that a college-age 354 who is 78 inches tall will normally weigh 210 pounds, but during training he will drop down to 190 or 185; he may make a good pitcher or first baseman. So would the 362, 363, or 364, but they would be more muscularly solid and powerful. In fact, the 362 would be seen far more often in big-league play than would the 262. To quote Sheldon, the 261, "together with its more numerous cousins the 263

and the 362, and its scarcer cousin the 462, defines the picture of the most encountered successful athletes. These four somatotypes make up about 4.8% of the male population."[56] Sheldon goes on to report that "even if we allow for a small sprinkling of 5s in mesomorphy among the top-flight athletes of the country, it is apparent that the four somatotypes named must constitute something in the neighborhood of two-thirds of them."

Whereas in baseball most of the players are 262s, 263s, 362s, or 462s, in professional football one finds more 7s in mesomorphy and extreme ectopenes (low in ectomorphy) like the 371, 471, and 561. In fact, the 372's such as Jack Dempsey, Gene Tunney, Red Grange, and Jim Thorpe are well known for their power. There are also a number of powerful mesomorphs high in endomorphy, such as the 551 or 561 who at 6 feet is a rare prize for "holding the line." This strong man could, at college age, train down from 232 to 210 pounds. The author is acquainted with one who played the line in college and worked weekends as a nightclub bouncer.

Choosing the Activity

Athletic ambitions are often frustrated simply because of unsuitable body type.

Encouraging the rather large endomorphic child, as the proud father sometimes does, to enter football merely because of his size makes little sense. The coach soon finds out where the difficulty is. From the athletic point of view, such somatotypes as the 225, 216, and 325 are definitely removed from it all. The 325, for example, knows better than to attempt to enter the activity. So does the extreme endomorph, low in the other components. But the 235 has just enough mesomorphy to enter athletic games and suffer frustration in not being able to achieve the level of most of the group.

Very much along the same line are the almost-balanced types, such as 334, 343, and 344 who seem to be neither fish nor fowl in their athletic achievement. They are often interested in physical activity but seem to be caught halfway between being an athlete and not being one. Sheldon finds in his delinquency studies that they try to be tough athletes or "regular fellows" but have some perplexing frustrations that occasionally lead to delinquency.

In general, some of the less extreme somatotypes need an understanding coach who will offer them some athletic guidance.

The Female Somatotype

It was originally thought that men and women expressed two different and more or less independent somatotype systems. It was therefore presumed that eventually a standardized and different set of objective criteria would be set up for women. Later, based on sex factor observations, Sheldon and his associates decided to hew to the single somatotype system.

[56] W. H. Sheldon, *op. cit.*, p. 119.

Women at all ages are much more endomorphic than men and heavier in proportion to stature. Where the mean ponderal index of 4,000 men was 13.19, the mean for 4,000 college women was only 12.84.[57] This could represent the endomorphic difference between a 335 and a 435. The commonest male somatotypes range between 344 and 353, whereas the commonest female types fall around 533. Where the male ectomorphic mesomorphs are plentiful, the female ectomorphic mesomorphs are considerably less frequent. It may appear, when the pilot study for *Atlas of Women* is completed, that a dozen or more of the male somatotypes are quite rare or not present in the

Fig. 13-17. The Mesomorphic Ectomorph Dancer (S-$3^2 44^2$) and the Mesomorphic Endomorph Field Hockey Player (S-$4^2 43$).

female population. For instance, the 1s in endomorphy and the 7s in mesomorphy seem to be missing.

It is altogether possible that because of sexuality, endocrine and metabolic differences induce a different expression of strength in the primary component. As it is now, two male somatotypes of 344 are about the same in motor ability potential, but a female 344 would be a powerhouse, the kind of body build generally seen among top ranking women athletes. It may be necessary, therefore, in order to compare the sexes, to set up a percentile distribution of the three components. In this way any somatotype can be translated from three components to three percentile scales for either sex.

Another point worth mentioning is the growth in male and female soma-

[57] *Ibid.*, p. 13.

totypes. Very often during the twelve- or eighteen-month period just preceding puberty, the boy or girl has what Sheldon calls a "fat period" or period of "endomorphic blossoming." This may be a preview of what the person will look like in later years and is especially noticeable with women. It makes the young female adolescent talk of reducing at a time when she may need more food intake just for energy to grow on. Gauging endomorphy in women, therefore, will require study and practice. An old photograph taken of the subject in a bathing suit before puberty should help.

The female g-index is an asset to motor ability and general physical performance. The broad shoulders and narrow hips of the male are conducive to body movement and sports skills. The mesomorphic ectomorph as a dancer or the ectomorphic mesomorph as a high hurdler represents structure in action, a kind of action simply impossible for the rotund viscerotonic to duplicate or even appreciate.

In an extensive statistical analysis of the motor ability of college women, Carruth found a high relationship with body coordination, a moderate relationship with power, strength, and speed, and slight positive relationship with flexibility, balance, endurance, and agility.[58] All these factors tended to depend upon the amount of the mesomorphic component. Ectomorphy seemed to be a limiting factor in effective motor performance, and all-round performance was better in the combination of mesomorphy and endomorphy. Morris discovered about the same result in her study of champion women athletes.[59] These athletes had a mean somatotype rating of $4^{2}43$; the nonathletes were 533.

Implications for Research

Human behavior is an intricate and thought-provoking subject. Athletic behavior, which is physical-social, depends to a great extent upon the qualities of each person.

Much research has been carried on in the past without proper respect for individual variation. In health and physical education, persons have been compared solely on the basis of sex, age, and weight. One thing that the constitutional researcher wants to do is lay a foundation of order and structure for the study of individual behavior differences.

Much of the educational, psychological, and medical-biological research needs to be carried out more specifically on well-identified persons. To say, as we do so often, that "1,000 males of college age were used in the sample" is far too general. It is possible that much of the so-called basic behavior research may need to be done over, for the results represent only a kind of

[58] Wincie Ann Carruth, "An Analysis of Motor Ability and Its Relationship to Constitutional Body Pattern of College Women," unpublished doctoral dissertation, New York University, New York, 1952.

[59] Patricia Collins Morris, "A Comparative Study of Physical Measures of Women Athletes and Unselected College Women," unpublished doctoral dissertation, Temple University, Philadelphia, 1960.

average somewhere between both ends of the scale of behavior, one extreme type of behavior canceling out the opposite type.

It should now be obvious to the reader that the constitutional make-up of boys and girls affords a rich area for additional research. Physical educators and nutritionists, as well as psychologists, actuaries, geneticists, and human biologists, stand to profit from this kind of study. Moreover, education, especially at this late date, needs a greater understanding of individual differences. As the sturdy oak and the rubbery willow differ in structure, so children of the same general appearance are typically different.

The best that can be hoped for in education is to improve the system within which teachers of average ability do their work. Constitutional analysis and application holds great promise. Finally, it fosters a profoundly democratic respect for the individual person. The philosophy of the well-worn poem of John Bangs expresses a closing point of view for this chapter. Bangs speaks of having met the little elf man "down where the lilies blow," and he asked him why he was so small, and why he did not grow.

> He slightly frowned and with his eye
> He looked me through and through—
> "I'm just as big for me," said he,
> "As you are big for you!"
>
> *The Little Elfman*

SELECTED REFERENCES

Bauer, J.: *Constitution and Disease*, 2d ed., New York, Grune & Stratton, Inc., 1945.

Bookwalter, Karl W.: "The Relationship of Body Size and Shape to Physical Performance," *The Research Quarterly*, 23:271–279, October, 1952.

Cureton, Thomas K.: *Physical Fitness Appraisal and Guidance*, St. Louis, The C. V. Mosby Company, 1947, chap. 4.

———: "Body Build as a Framework of Reference for Interpreting Physical Fitness and Athletic Performance," *The Research Quarterly* (Supplement), 12:301–330, May, 1941.

Damon, A.: "Physique and Success in Military Flying," *American Journal of Physical Anthropology*, 13:217–219, June, 1955.

DiGiovanna, Vincent: "The Relation of Selected Structural and Functional Measures to Success in College Athletics," *The Research Quarterly*, 14:2–6, May, 1943.

Draper, George, Clarence W. Dupertius, and J. L. Gaugher, Jr.: *Human Constitution in Clinical Medicine*, New York, Paul B. Hoeber, Inc., 1944.

Feigenbaum, J.: "Physical Constitution and Disease," *Archives of Internal Medicine*, 55:445–456, March, 1935.

——— and D. Howat: "The Relation between Clinical Constitution and the Incidence of Disease," *Journal of Clinical Investigation*, 13:121–138, January, 1934.

CHAPTER 14

Classification Measures

Educational classification is involved in managing persons. It is the basis on which school children are grouped for an activity, and it is directly related to a knowledge of their primary abilities Logically, therefore, it would be most difficult to give full recognition to the educational significance of individual differences, needs, and capacities without some system of classification.

Purposes of Classification

From an administrative point of view it is frequently desirable, if not downright necessary, to classify pupils in physical education This is particularly true in large secondary schools, where the program has to be adjusted to large numbers of pupils

As a rule the chief purpose behind the classification of school children is to equate boys and girls according to their respective abilities and to facilitate efficient teaching Thus pupils with low physical fitness scores may be grouped for individual remedial instruction, pupils rated in athletic ability may be grouped for sports participation, and students possessing similar sensory defects may be placed in special classes or receive a modified type of instruction

Although there are arguments both for and against homogeneous grouping, it must be admitted that it is immediately more efficient in physical education to work with pupils of similar abilities than to struggle with a heterogeneous mixture of age, maturation, physique, skill, knowledge, interest, and physical fitness Such heterogeneity creates obstacles to advancement. Major criticism of many physical education classes is that there is little progression from one year to the next and no differentiation within the same class of varying pupil abilities

In a number of secondary schools simple sports skill tests have been used to classify pupils for instruction, with the result that the pupils have not only enjoyed the sport but gained a rich experience in skills and knowledge. In Arlington, Virginia, for example, the classification test scores are tabulated and graphed so every class can be divided each into three or four groups according to the results[1] There is full understanding that pupils can be

[1] See a somewhat classic example of this by Julian U Stein, "Better Basketball through Skill Classification," *Journal of Health, Phyical Education and Recreation,* 28:10–11, November, 1957.

moved up or down out of groups during the course of the instructional periods if they progress more or less rapidly than the remainder of the group.

In college and high schools motor ability measures are frequently used to classify students for instruction. In colleges and universities it is possible to give partial credit for skill and knowledge in a physical education area in order to move a student to a category in which he shows inexperience or a general weakness. Lockhart and Mott used the Scott Motor Ability Test to classify college women and found that the great majority of students preferred classes limited to persons of similar ability.[2]

In physical education, practically every test given can be used for classification purposes. Some are used for diagnostic reasons, others as a basis for setting up intramural team competition, and still others for sectioning students in required or adapted physical education classes. In one school, for instance, there are four graduated levels of classification for physical education:

1. Unlimited physical education including interscholastic sports.
2. Restricted physical activity, limited to classwork and excluding interscholastic and intramural strenuous sports.
3. Adapted individual physical education, for pupils with heart conditions, postoperative and postillness cases, children with postural defects, etc.
4. Social recreation, including quiet games involving little or no physical exertion.

Means of Classification

There are a number of ways to classify pupils of any age for physical education activities. One of the oldest and most popular is the age-height-weight method. There is some logic in this method because children, in general, gain in height and weight as they grow older, and this is usually accompanied by increased strength, power, and motor coordination. Above the age of seventeen, however, age, height, and weight are not too satisfactory as bases for norming athletic performance.[3] Other factors are of much greater significance, such as muscle strength, specific skill experiences, and general coordination.

Another weakness in the age-height-weight method is that it has little respect for maturity or body build. A large eighth-grade endormorph is entirely different from a mesomorph of essentially the same height, weight, and chronological age. Moreover, it is likely that he is far less mature

[2] Aileene Lockhart and Jane A. Mott, "An Experiment in Homogeneous Grouping and Its Effect on Achievement in Sports Fundamentals," *The Research Quarterly*, 22:58–62, March, 1951.

[3] Thomas K. Cureton, "Body Build as a Framework of Reference for Interpreting Physical Fitness and Athletic Performance," *The Research Quarterly* (Supplement), 12:314–320, May, 1941.

physiologically. Individual maturation, especially in the elementary–junior high school period, is a significant item.

Two indices long used in equating teams for athletic competition are McCloy's Classification Index and the Neilson and Cozens Classification Index.

McCloy's Classification Index. By weighting age, height, and weight McCloy built an index that correlated well with track and field events and fairly

Classification index	Class	For small groups	Class	For large groups
I—high school	A	890 and over	A	900 and over
	B	860	B	845
	C	830	C	815
	D	800	D	785
	E	770	E	755
	F	740	F	725
	G	739 and under	G	685
			H	665
			I	664 and under
II—college men	A	570 and over	A	580 and over
	B	550	B	560
	C	530	C	540
	D	529 and under	D	520
			E	519 and under
III—elementary school	A	275 and over	A	275 and over
	B	260	B	263
	C	245	C	255
	D	230	D	245
	E	215	E	235
	F	200	F	225
	G	185	G	215
	H	184 and under	H	205
			I	195
			J	185
			K	184 and under

Fig. 14-1. McCloy's Classification Index Divisions.

well with selected sports skills.[4] There are actually three indices: one for high school, where height appears important; one for college, where age above seventeen is negligible; and one for the elementary grades, where height does not seem to be too significant:

High school: Classification index I = (20 × age) + (6 × height) + weight
College men: Classification index II = (6 × height) + weight
Elementary school: Classification index III = (10 × age) + weight

[4] Charles McCloy, *The Measurement of Athletic Power*, New York; A. S. Barnes and Company, 1932.

Once the classification index for a student has been determined, it is related to another classification whereby the student is grouped according to Class A, B, C, D, etc., as in Figure 14-1.[5]

To illustrate, if a high school boy was to be classified for competition in a large group activity, such as soccer, and he was fifteen years of age and 60 inches tall and weighed 110 pounds, his classification index, $(20 \times 15) + (6 \times 60) + 110$, would equal 770. He would be placed in the E group for activity. As indicated elsewhere in the text, the classification index is used with the Carpenter Motor Ability Test and the Indiana Physical Fitness Test.

Neilson and Cozens Classification Index. This index, developed in California about 1930, has had wide use in upper elementary and secondary schools.[6] It is based on the exponent system, where age, height, and weight

Exponent	Age (in years and months)	Height (in inches)	Weight (in pounds)	Sum of exponents	Cla
1	10-10.5	50-51	60-65	9 and below	A
2	10.6-10.11	52-53	66-70	10-14	B
3	11-11.5		71-75	15-19	C
4	11.6-11.11	54-55	76-80	20-24	D
5	12-12.5		81-85	25-29	E
6	12.6-12.11	56-57	86-90	30-34	F
7	13-13.5		91-95	35-38	G
8	13.6-14.11	58-59	96-100		
9	14-14.5		101-105		
10	14.6-14.11	60-61	106-110		
11	15-15.5		111-115		
12	15.6-15.11	62-63	116-120		

Fig. 14-2. Elementary School Boys, and Girls, Classification Chart (Neilson and Cozens).

are related to a number instead of to a letter, as in McCloy's index. In California it is referred to as the California plan of classification and is used with the California Physical Performance Test (see Chapter 8). The exponent for each factor (age, height, and weight, is set forth in Figure 14-2, and the sum of the three exponents is the key to the classification. Exponents have been set up for boys and girls in grades five through eight and for boys in junior and senior high school (see Figures 14-2 and 14-3). An example of how the Neilson and Cozens Classification Index works may be taken from Haverford Junior High School, Haverford, Pennsylvania. This

[5] Charles H. McCloy, *Tests and Measurements in Health and Physical Education,* New York, Appleton-Century-Crofts, Inc., 1935.

[6] N. P. Neilson and F. W. Cozens, *Achievement Scales in Physical Education Activities for Boys and Girls in Elementary and Junior High Schools,* Sacramento, Calif., California State Department of Education, 1934.

school of over 1,600 pupils plays regularly scheduled interscholastic football under strict regulations controlling the coaches and players. Members of the school team must have a total exponent of 20 or less. Baseball pitchers cannot be over 22. Also, in the nine elementary schools, instead of a chosen few playing interscholastic sports, the opportunity is open to all boys and girls. In a typical year, 62 per cent of the sixth grade boys participated in

Exponent	Age (in years and months)	Height (in inches)	Weight (in pounds)	Sum of exponents	Class
12	....	....	72–78	88 and over	A
13	....	....	79–84	83–87	B
14	....	....	85–90	79–82	C
15	....	....	91–96	75–78	D
16	....	....	97–103	70–74	E
17	....	....	104–109	69 and below	F
18	....	....	110–115		
19	....	....	116–121		
20	....	....	122–128		
21	....	....	129–134		
22	10.9–11.2	47 and less	135–140		
23	11.3–11.8	47½–49	141–146		
24	11.9–12.2	49½–51½	147–153		
25	12.3–12.8	52–53½	154–159		
26	12.9–13.2	54–55½	160–165		
27	13.3–13.8	56–57½	166–171		
28	13.9–14.2	58–59½	172–178		
29	14.3–14.8	60–62	179–184		
30	14.9–15.2	62½–64	185–190		
31	15.3–15.8	64½–66	191 and over		
32	15.9–16.2	66½–68			
33	16.3–16.8	68½–70½			
34	16.9–17.2	71–72½			
35	17.3–17.8	73–74½			
36	17.9–18.2	75 and over			
37	18.3–18.8				
38	18.9–19.2				

Fig. 14-3. Junior-Senior High School Classification Test (Neilson and Cozens).

league football games, and 80 per cent of the sixth grade girls played inter-school hockey. Exponents are used to set up competitive classes.

At the college level, Miller demonstrated that the Neilson and Cozens index was quite satisfactory for equating groups of men according to body size.[7] The idea of classifying men for physical activities according to physique has merit and needs more experimentation to give it greater support.

[7] Kenneth T. Miller, "A Critique on the Use of Height-Weight Factors in the Performance Classification of College Men," *The Research Quarterly*, 23:402–416, December, 1952.

Earlier, Miller found the Wetzel Grid to be a satisfactory tool in classifying college men by physical make-up.[8] The grid, however, was weak in reliably classifying motor performance. On the other hand, Sills and Mitchem found that a classification of body build into thirteen somatotype catagories correlated with three tests of physical fitness quite well.[9] Multiple correlation coefficients for body build with sit-ups, pull-ups, and the 300-yard shuttle run showed substantial relationships that could serve as a basis for classifying male college students into homogeneous groups, a process that makes it possible to grade students against others of similar ability.

Equalization of Teams

In addition to classification indices and motor performance tests there is another method which has been employed successfully for some time. This is the strength index (SI).

Like other measures used for equating teams, the SI is based on the *doctrine of equality*. It supports the remark of Cicero, "Put yourself on a level with your friend." To do this is democratic activity, for when the teams and players are equal, there is in fact a very *just* situation. Whenever one side has greater power, the situation is not just There is little opportunity to develop social efficiency through sportsmanship, the essence of which is fair play. Hence, justice, which philosophers find hard to define, is understood by the sportsman to mean "no unfair advantage," all of which suggests that *sportsmen, or persons to be trained in sportsmanship, should be equal in strength, skill, and intelligence.*

Of course, this is hard to do, but the dedicated physical educator can certainly try. One begins by refining the measure of choosing teams. This brings up the topic of strength and power, which cuts across age and maturation levels and is related positively to athletic ability.[10] A certain minimum of strength is necessary to carry out specific game skills. In grouping children where the amount of specific skill training is frequently about equal, a group of pupils selected at random will tend to perform a given test according to their index of strength. That is, those pupils who are stronger will run and jump and achieve better results in many isolated skills than those pupils weaker in physical strength.

The SI has been used successfully for a number of years chiefly to equate intramural teams. Much of this has been done in New York State, Oregon, and the New England states. In Figure 14-4 there is an example of how the

[8] Kenneth T. Miller, "The Wetzel Grid as a Performance Classifier with College Men," *The Research Quarterly*, 22:63–70, March, 1951.

[9] Frank D. Sills and John Mitchem, "Predicting Performances on Physical Fitness Tests by Means of Somatotype Ratings," *The Research Quarterly*, 28:65–71, March, 1957.

[10] Carl E. Willgoose, "The Relationship of Muscular Strength to Motor Coordination in the Adolescent Period," *Journal of Educational Research*, 44:138–143, October, 1950.

strength index can be used to create divisions for the classification of boys for participation in a Youth Week track meet.[11] The scores represent performance of secondary schools boys in Andover, Massachusetts.

Strength index	Division	Winner's record (in feet and inches)	Number of participants	Median performance (in feet and inches)
Running broad jump:				
Under 950	1	12, 10½	31	10, 2
950–1149	2	13, 1	20	11, 6
1150–1349	3	15, 0	18	12, 11
1350–1549	4	15, 4	12	14, 10
1550–1749	5	15, 8	11	15, 5
1750 up	6	16, 11	17	15, 5
Running high jump:				
Under 950	1	4, 1	10	3, 8
950–1149	2	4, 2	15	3, 10
1150–1349	3	4, 7	20	4, 2
1350–1549	4	4, 8½	19	4, 2
1550–1749	5	5, 1	9	4, 7
1750 up	6	5, 2	7	4, 8
8-lb shot-put:				
1550–1749	5	37, 5	10	33, 3
1750 up	6	44, 9	8	39, 3

Fig. 14-4. Classification by the Strength Index.

One of the most admirable studies done on the use of the strength index in the equalization of teams was done by Leonard Clark, in Melrose, Massachusetts.[12] Of fifty-seven intramural football games, 35 per cent ended in tie scores. Of seventy-two ice hockey games played, over 60 per cent were either tied or won by 1 point. Yavits obtained similar results at Ithaca College with male students.[13]

Where pupils are classified according to similar abilities, and they know this, there is considerably more interest in team play. The number of participants in an intramural activity increases. There are more spectators as news of close contests spreads through the community. Moreover, parents endorse the program and frequently join with the administrator in expressing the view that "a tie score is an ideal score," a view akin to the truth when pupils are properly classified.

[11] Carl E. Willgoose, "Use of Strength Tests in Team Equalization," *The Physical Educator*, 6:4–6, May, 1949.

[12] Leonard Clark, "Athletics and Sportsmanship: Policies and Techniques," *Education*, 60:517–520, April, 1940.

[13] Isadore Yavits, "Equalizing Opponents in a Teacher Training School," *Education*, 60:524–525, April, 1940.

GRADING IN HEALTH AND PHYSICAL EDUCATION

When it comes to marking pupils on their achievements, health education presents no problem. But in physical education there is a wide range of opinion on why and how to grade.

Fundamentally, some kind of grade is necessary in every school subject to show the degree of progress to the students and parents. Although the evils of the marking system are well known, marks, nevertheless, have become symbolic of school achievement.[14] Thus, when they are not given in physical education, both parents and students tend to relegate physical education to the level of an unimportant subject. Although this is far from the truth, the impression is quite the opposite.

The need for meaningful marks based on specific accomplishments is ever-present. The relationship here between proper appraisal techniques and status reporting is clear. The two go together. The keener the evaluation, the more valuable the grade.

Grades, and satisfactory techniques for determining them. are only half the problem. The other half is credit. Just as pupils receive marks in English, social studies, science, etc., they also receive credit. The numerical credit in the secondary schools, like the mark itself, gives status to the activity. Of course, physical education has progressed a long way on its own merit and without the aid of recognized academic credit; but it still occupies a different footing than other subjects. Moreover, meeting a mandatory time requirement, such as 300 minutes a week in New York State, is not appreciated by school personnel and pupils as much as definite units of credit.

There is a strong possibility that giving marks and credit will encourage evaluation practices within the profession. Since evaluation involves goals and program content, these factors should also be advanced. Many large school systems in several states have set up specific criteria for determining progress marks.

Grading large numbers of pupils may take time in physical education, especially where large classes are common and grades are sent home several times during a school year. But it is worth the effort, for a piece of tangible evidence of a child's status, based on objective appraisals, can be a most informative report to parents and others.

Methods of Grading

Unfortunately, one of the most common ways of grading in physical education is "pass" or "fail"; this is frequently expressed another way: S for satisfactory and U for unsatisfactory. Experience indicates that in this method of grading practically everyone gets a passing mark, which in itself degrades the whole process. The mark means very little. It cannot discrimi-

[14] For a complete coverage of the weaknesses in marks, see reference to Clay C. Ross at end of chapter.

nate quality performance, and all too often it represents a grade based on such weak items as gymnasium uniforms and attendance.

Another way to mark pupils is to employ numerical grades such as 5, 4, 3, 2, and 1 to represent whole letter grades of A, B, C, D, and E, respectively. Also in this manner letter grades can be assigned to various physical education items or areas. These areas can be weighted by the instructor to indicate greater importance than other areas of study. For example, the level of a student's physical fitness might be weighted 4 points, motor skills 3 points, social efficiency 2 points, intramural participation 1 point, and uniform and appearance 1 point. When a student is marked in each of these areas from A to E (that is, from 5 to 1 numerical points), the result is multiplied by the weighting figure (4, 3, 2, or 1) to produce a final score or grade, as follows:

Grading item	Weight factor	Mark	Points
Physical fitness	4	A(5)	20
Motor skills	3	B(4)	12
Social efficiency	2	B(4)	8
Intramural participation	1	C(3)	3
Body mechanics and appearance	1	B(4)	4
Total score or grade score			47

To convert the raw score of 47 into a single percentage grade for school reporting purposes, one might divide it by the best possible score, which is 55 points, and multiply it by 100. The result is 85 per cent. In this instance the pupil knows that he lacks only 15 per cent from having achieved a perfect mark. If necessary, this figure can be changed to a letter grade by setting up categories:

A is the highest 5 to 10 per cent.
B is the next highest 20 to 30 per cent.
C is the middle 35 to 45 per cent.
D is the next lowest 20 to 30 per cent.
E is the lowest 5 to 10 per cent.

It should be pointed out here that in the final analysis grading levels or categories are somewhat arbitrary. Where the "pass" line is drawn depends upon the individual teacher and local circumstances. One teacher, for example, may decide to split the distribution of scores into three pieces and rate everyone as good, poor, and fair. Very often, if the distribution of grade scores is fairly well balanced with high and low marks, the center scores will be worth looking at. Here the average score, mid-point score (median), and most typical score (mode) would fall at about the same point. In such case this location could be considered average or fair for the group tested. A liberal arrangement would be to have the C grade fall 20 points above and

below the mid-point of the distribution. The other grades could be set up as follows:

A	91–100 points	10 per cent of scores
B	71–90 points	20 per cent of scores
C	31–70 points	40 per cent of scores
D	11–30 points	20 per cent of scores
E	0–10 points	10 per cent of scores

The technique for transmuting point scores into letter marks should be sufficiently flexible to accommodate itself to differences that occur in groups of scores in central tendency and variability. When all the scores are fairly close together, or they are bunched in the middle of the distribution, or there

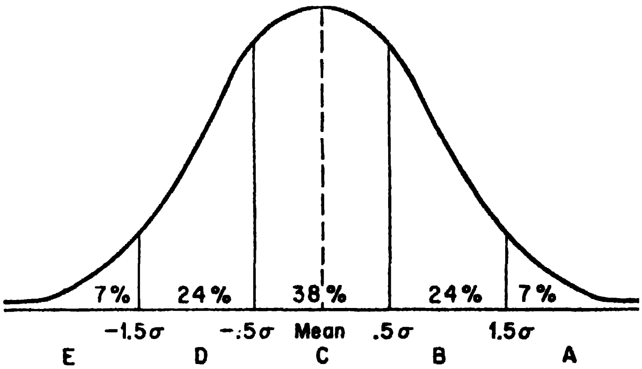


Fig. 14-5. Five-letter Marking System Based on the Normal Distribution Curve.

are about as many high scores as there are low, the above procedure works fairly well. When, however, there is a great amount of variability of scores about the center, another technique of grading may prove more accurate and more fair to the pupils. Here the mean (M) is used as a measure of central tendency, and either the standard deviation (σ) or the mean deviation (MD) is used as a measure of variability.

In a five-letter grading system based upon M and σ, the various letters are set forth as follows:

A = 1.5σ or more above the mean score
B = between $.5\sigma$ and 1.5σ
C = between $-.5\sigma$ and $+.5\sigma$
D = between -5σ and -1.5σ
E = 1.5σ or more below the mean score

In the normal curve of distribution of scores, the various letter grades would appear as in Figure 14-5. To illustrate how this might be used, the following situation is presented:

Example

Archery examination: 100 archery skill tests.
Range of scores: 30 to 80
Mean score = 54
Standard deviation(σ) = 6
Letter grades:
A = 66 and up
B = 58–65
C = 51–57
D = 43–50
E = 42 and below

When mathematics is involved, as has been the case up to now, the objectivity of the test mark is usually advanced over a purely subjective appraisal. Yet it is perfectly possible for the instructor to rate boys and girls with the aid of a check list or rating scale and report the findings in writing to the parents. Such a descriptive technique can be quite informative, provided the instructor takes the time to use well-chosen words and sentences for each student. As might be expected, this is especially difficult to do on a large number of pupils, particularly those in the middle or fair-to-good range. This is another way of saying that it is easy to appraise the good students and describe the difficulties of the poor student but hard to write about the average student who minds his own business and makes a satisfactory performance.

Reporting the Grades

By now it should be clear to the student of evaluation in health and physical education that marking is concerned chiefly with specific objectives and their accomplishment. School marks, therefore, to be meaningful should not be limited but should be related to the broad and varied program of activities at any grade level. A grade should tell the parent that the child is generally good or generally poor; it should also tell where the strengths and weaknesses are. This is particularly true in physical education.

To do this right away in the fall of the school year the instructor will have to do more than appraise a fall sport or early level of physical fitness. He will also have to refer to knowledge, social competency and attitude, posture, and such other items as attendance, uniform, and cleanliness. The amount of weight given any one item will vary. Some secondary schools have a marking system similar to the one used at the University of Florida. Here the student is evaluated on four points: skill, knowledge, fitness, and evaluation of the student by the instructor. In the latter case the instructor rates the student using a special instructor analysis sheet, the essence of which includes the following:

Instructor Analysis of Student

Attendance:

5—Excellent:	Student has perfect attendance and was never late for class.
4—Good:	Student has 1 absence and/or 1 or 2 lates.
3—Average:	Student has 2 absences and/or 1 or 2 lates.
2—Below average:	Student has 3 absences and/or 1 or 2 lates.
1—Poor:	Student has 4 absences and/or 1 or 2 lates.

Note: Excused absences are to be evaluated by the instructor.

Improvement: An improving, especially betterment; state of being improved; especially enhanced value or excellence.

5—Excellent:	Student has approximated the limit of his ability.
4—Good:	Student has shown improvement in all areas of the course.
3—Average:	Student has shown improvement in certain areas of the course.
2—Below average:	Student has shown improvement but not in line with his potential.
1—Poor:	Student has shown very little improvement.

Participation: To have a share in common with others, to partake, to share.

5—Excellent:	Student has at all times shown an eagerness to participate in the activities of the course.
4—Good:	Student has lapsed in his participation at times during the course.
3—Average:	Student has participated in the activities of the course under motivation.
2—Below average:	Student participated, but at times he had to be motivated by special request of the instructor.
1—Poor:	Student had little desire to participate and had to be continually reminded by the instructor that activity was a part of the course.

Attitude: Position or bearing as indicating action, feeling, or mood.

5—Excellent:	Student was at all times a highly desirable member of the class.
4—Good:	Student had desirable habits of group interaction.
3—Average:	Student was a desirable member of the class, but at times he showed minor undesirable traits of action.
2—Below average:	Student could be singled out as having social traits that were undesirable.
1—Poor:	Student had traits that definitely needed counseling for purposes of correction.

This form can also be used for constructive self-analysis. Experience at the University of Florida indicates that self-knowledge presents an excellent opportunity to the student for character development and self-realization. Thus students are frequently asked to rate themselves early in the term, at mid-term, and at the end. The raw numerical scores from these instructor analysis forms are converted to T scores and added to the T scores on skills, knowledge, and physical fitness for a combined grade. This total T score is then converted to a letter grade and is the student's mark for physical education.

Because progress reports in physical education are so important to the parents, Miller has developed a form letter designed to inform parents about the marking system.[15] It precedes the arrival home of the actual grades (see Figure 14-6).

[15] Copies may be obtained from Arthur G. Miller, Boston University, Boston. Mass.

SUGGESTED FORM LETTER TO PARENTS
(Marking)

Dear ____________________ (Parent)

This letter is intended to acquaint you, and to give you information concerning our new form for reporting your child's achievement in the physical education program here at ____________________ (School).

As indicated on the form, two reports will be issued yearly -- in January and June. Three areas of physical education are marked. These three areas are the physical, social and intellectual objectives of our physical education program. A five point rating scale of superior, good, average, poor and inferior is used in all three areas. All the activities in our physical education program are not necessarily included among the skills listed. The activities listed are those which we consider core activities in our program.

Achievement is reported as follows:

1. Physical

   Fitness - by performance in areas such as strength, endurance, agility and balance. This test is not the same for boys and girls.

   Skills - by testing sports skills in the program. Some apply to girls only (field hockey), some to boys only (football) while others apply to both boys and girls (basketball). Skills test are not the same for boys and girls.

2. Social

   By observations of the instructor in those traits listed. These observations occur in class, game, team and individual activities. An example of leadership would be the ability to act as a squad leader and perform his duties and functions well. Health habits include cleanliness, appearance, shower procedure and uniform cleanliness.

3. Intellectual

   By a test of the pupil's knowledge of fundamentals, history, rules and strategy of the sports activities.

We hope that this type of report will meet with your approval, and that it will aid you in better understanding the progress your boy or girl is achieving in our physical education program.

Any suggestions or remarks you may have concerning this report will be greatly appreciated and may be included on the form when it is returned.

Sincerely yours,

Director, Physical Education

Fig. 14-6. Orientation Letter to Parents on Marking. (Courtesy Arthur G. Miller.)

As noted in the orientation letter, grades are suggested for twice a year, such as September to January and January to June This allows the instructor more time and permits a more thorough report At these times a physical education achievement record is sent to the parents The first

First Semester (September January)

Year

	Superior	Good	Average	Poor	Inferior
PHYSICAL					
Physical Fitness					
Sports Skills					
Soccer					
Field Hockey (G)					
Football					
Speedball					
Basketball					
PHYSICAL AVERAGE					
SOCIAL					
Leadership					
Interest					
Cooperation					
Health Habits					
SOCIAL AVERAGE					
INTELLECTUAL					
Sports Skills					
Knowledge Test					
INTELLECTUAL AVERAGE					
COMMENTS					

Fig 14-7 Physical Education Achievement Record–First Semester (Courtesy Arthur G Miller)

semester report averages the physical, social, and intellectual components, respectively (see Figure 14-7)

One other method of reporting marks in physical education to parents that merits more attention than it gets in most places is the parent-teacher conference Here the parent or parents come to the school to talk with the

instructor of physical education. This method of reporting provides an opportunity to discuss low fitness, poor health habits, and the strong and weak points of the pupil. It is helpful to the instructor to have on hand at the time of the parent-teacher conference a combined pupil record covering the achievements, abilities, and attitudes to be discussed. An excellent form for such a purpose has been devised for use at the Craig School, Schenectady, New York.[16]

Summarizing this chapter, it can be said without hesitancy that classification of pupils is a means to an end. The end is better instructional methods and improved grading practices, both of which contribute measurably to educational progress in health and physical education and a keener understanding of this field by school administrators and taxpayers.

SELECTED REFERENCES

Adams, Eleanore E.: "The Study of Age, Height, Weight, and Power as Classification for Junior High School Girls," *The Research Quarterly,* 5:95–99, May, 1934.

Bookwalter, Karl: "Marking in Physical Education," *Journal of Health and Physical Education,* 7:22–24, January, 1936.

Broer, Marion R.: "Are Our Physical Education Grades Fair?" *Journal of Health, Physical Education and Recreation,* 30:27-28, March, 1959.

Cox, Walter A., and Kenneth B. DuBois: "The Strength Index in Equating Intramural Teams in Albany, New York," *The Research Quarterly* (Supplement), 6:202–206, March, 1935.

Fairfield, Paul: "Credit for Physical Education," *New York State Journal of Health, Physical Education and Recreation,* 10:36–38, June, 1958.

Haag, Jesse H.: *School Health Program,* New York, Holt, Rinehart and Winston, Inc., 1958 p. 498.

Handy, Donald T., and Marjorie Latchaw: "Value of Academic Grades in Estimating Student Teaching Success," *The Research Quarterly,* 28:347–351, December, 1957.

Hinton, Evelyn A., and Lawrence Rarick: "The Correlation of Rogers Test of Physical Capacity and the Cubberly and Cozens Measurement of Achievement in Basketball," *The Research Quarterly,* 11:58–64, October, 1940.

Kistler, Joy W.: "A Comparative Study of Methods for Classifying Pupils," *The Research Quarterly,* 5:42–45, March, 1934.

McCall, William A.: *Measurement,* New York, The Macmillan Company, 1939, book 6.

Oberteuffer, Delbert: *Physical Education,* New York, Harper & Brothers, 1951, pp. 133–134.

Ross, Clay C.: *Measurement in Today's Schools,* 3d ed., Englewood Cliffs, N.J., Prentice-Hall, Inc., 1954, chap. 14.

[16] Copies can be obtained from J. E. Torpey, Department of Physical Education, Craig School, 2566 Balltown Rd., Schenectady, N.Y.

CHAPTER 15

Measurement of Social Efficiency

One of the professed aims of physical education is to teach young persons how to get along with others and exhibit desirable standards of conduct. This is true in health education as well, for social health, according to the World Health Organization, is on a par with physical, mental, and spiritual health. Moreover, social efficiency is an aim not foreign to any area of education. The degree to which students become socially adjusted depends not alone on any one field of education but on the combined efforts of all school personnel.

In a very real sense the heart of education is social efficiency. For years it was referred to as "character education." Only recently has it been spoken of as *human relationships* and *civic responsibility,* where youth and adults interact harmoniously in the local community, the state, the country, and the brotherhood of nations.[1]

It is most logical that health education teachers in the classroom and physical education instructors on the playground and gymnasium observe, and concern themselves with, the social behavior of their pupils. The seriousness of their concern is known through some effort to appraise social status, within the limits of their technical skill. This appraisal effort is purposely qualified, and there is a good reason for this. First of all, behavior is complex, and social behavior is of particular concern to the school psychologists, personnel workers, and guidance specialists. These persons have a special training for this work not generally possessed by health and physical education teachers, and they are the best qualified to evaluate social behavior, make diagnoses, and seek causes and effects. This is a service about which physical education personnel should have some knowledge and understanding, but the technical aspects of precise measurement and the strengths and weaknesses of current social adjustment batteries demand a specialist. Thus the psychologist or guidance worker is a source to be called upon to help the physical instructor or athletic coach determine the cause of poor adjustment. Such cooperative effort is most productive in many schools today.

Promoting Social Efficiency

Health and physical education personnel have long recognized their unique position in fostering citizenship and the social graces. Because boys and girls generally like physical activity, games, and play, they frequently look upon their instructors and coaches more as immediate friends than

[1] See reference to the Educational Policies Commission in Chap. 2.

typical teachers. It has been demonstrated on a number of occasions that students *do* come to their physical education teachers with a great variety of problems and that a large percentage of physical education teachers carry responsibilities in connection with the more formal guidance activities of the school.[2] Moreover, a large number of physical education teachers are called upon to confer with classroom teachers and serve as members of individual case conferences.

Stevenson, medical director of the National Association for Mental Health, has said that physical education has an advantage in promoting social well-being because it is incompletely identified with traditional education and can capture the response of the child by satisfying present needs more than any other aspect of education.[3] This viewpoint is shared by Bonney and Burleson, who stress the value of socializing techniques and the sociometric grouping of pupils in physical education classes.[4] Also it is quite probable that the sociometric organization of classes is most profitable at the elementary school level.

Social Efficiency and Physical Education

Social efficiency can be measured. It is made known through participation in group games, intramural and interscholastic sports, dance, and numerous other physical education situations that call for a demonstration of cooperation, leadership, initiative, perseverance, loyalty, courage, friendliness, honesty, and justice. Moreover, there is abundant evidence that these traits carry over to other youth and adult activities. In fact, the primary elements of motor performance tend to be related to certain character traits of leadership and popularity.[5] Here we have physical fitness and motor ability, especially during the adolescent period, providing among other things a *capacity* to understand others. This capacity enhances prestige and leads to greater social acceptance.[6] It frequently comes out in personality studies. In Sperling's study, for example, five separate personality inventories indicated that varsity and intramural participants scored significantly higher on personality adjustment, ascendence, and extroversion than did nonathletes.[7] In Walters' study of bowlers, a comparison of social adjustment ratings between

[2] Eloise M. Jaeger and Helen M. Slocum, "Physical Education Teachers Contributions to Guidance in Minnesota Secondary Schools," *The Research Quarterly*, 27:23–31, March, 1956.

[3] George S. Stevenson, M.D., *Children in Focus*, National Education Association and American Asociation for Health, Physical Education and Recreation, Washington, 1954, p. 185.

[4] Warren C. Bonney and Reba M. Burleson, "Socializing Techniques," *Journal of Health, Physical Education and Recreation*, 25:40–42, March, 1954.

[5] See particularly the reference to Jones, Cowell, Fraleigh, McCraw and Tolbert, Fulton and Prange, and Breck at the end of the chapter.

[6] Raymond G. Kuhlen and Beatrice J. Lee, "Personality Characteristics and Social Acceptability in Adolescence," *Journal of Educational Research*, 34:201–208, September, 1943.

[7] H. P. Sperling, "Relationship between Personality Adjustment and Achievement in Physical Education Activities," *The Research Quarterly*, 13:191–197, October, 1942.

good and poor bowlers showed the good bowlers to be the better accepted members of the group.[8] In a study by Biddulph involving 461 sophomore and junior high school boys, it was found that the personal and social adjustment of boys with high athletic achievement was significantly greater than the adjustment of boys ranking low in athletic achievement.[9] Walters recommended that boys with low motor ability need to be better understood, especially in the home, and that parents should be informed of the relationship of athletic achievement to social adjustment so they will appreciate the adjustment problems that may be associated with a lack of motor skill.

The personality of an individual and his efficiency as he socializes with others in a complex society is difficult to simplify. Individual attitudes are involved, which are frequently hidden. Fortunately, through the overt expression peculiar to vigorous physical activity, many attitudes are made known directly to the instructor. This, coupled with an attitude questionnaire, may give a fairly accurate picture. Attitude instruments, however, are among the most difficult to build. The better ones usually describe a situation and ask the student to comment on it by checking one of several possible choices. McAfee set up twenty such situations for sixth-, seventh-, and eighth-grade boys to rate on a test designed to measure sportsmanship.[10]

Equally as difficult to deal with are the moral and spiritual values in education. Yet physical education, particularly through group sports, provides the opportunity to develop sound moral values by playing according to the rules (even when the official is not looking), by respecting authority, by exercising fair play, by being thoroughly honest, and by reaping a genuine feeling of satisfaction after a hard-played activity. Moreover, this can be improved upon to some extent by permitting the athletes themselves to govern and direct the sports event. By this is meant that at game time the coach sits up in the bleachers and permits team managers and captains to direct their own activity and gain the rich reward of applying what they have been taught without the immediate comments of their coach. This has been referred to as *player control,* an educational concept developed in the early 1930s by Frederick Rand Rogers but not given a wide trial because of the general reluctance of coaches to take a chance on losing a game in a situation where there is pressure to win. If the practice was properly tried out, possibly in a league, with the mutual consent of all teams, its values could be properly appraised and made known. Potentially it should do more for leadership, self-realization, and team cohesiveness than other schemes to encourage social development.

Certainly physical education affords a realistic laboratory experience in democratic living, and a large share of this can be realistically evaluated.

[8] C. Etta Walters, "A Sociometric Study of Motivated and Non-motivated Bowling Groups," *The Research Quarterly,* 26:107–112, March, 1955.

[9] Howell G. Biddulph, "Athletic Achievement and the Personal and Social Adjustment of High School Boys," *The Research Quarterly,* 25:1–8, March, 1954.

[10] Robert H. McAfee, "Sportsmanship Attitudes of Sixth, Seventh, and Eighth Grade Boys," *The Research Quarterly,* 26:120, March, 1955.

Appraising Social Efficiency

A number of instruments that attempt to measure some phase of personality, mental and emotional adjustment, or social well-being have appeared over the years. These vary from simple observation scales to more elaborate and highly technical diagnostic tools.

In physical education, McCloy was one of the first to point to the need for appraisal in this area. His work on character building excels as foundation reading for the dedicated student.[11]

OBSERVATION AND RATING SCALES

In order to make simple day-to-day observations more objective it has been necessary to devise check lists and rating scales. Here the characteristics of proper social behavior are generally listed so they can be appraised for degree and frequency. For example, a typical inquiry might be how often student Jones has shown signs of leadership and to what extent.

In all grades certain signs or patterns of behavior are evidence of good social and emotional health: when the pupil seems happy most of the time, joins others for play and enjoys the group, is chosen by others for study groups, committees, and games, does not daydream excessively, exercises adequate self-control for his age, and is able to make and maintain friends.

In the lower grades the teacher can fashion a scale where social characteristics are rated as "very good," "good," "poor," and "very poor." "Very good" may represent the top 10 per cent of the class; "good," the group above average; "fair," the typical group; "poor," the below-average group; "very poor," the bottom 10 per cent of the class. To this can be added the "undecided" or "don't know" category reserved for pupils about whom there is insufficient evidence to permit judgment. These qualities can be scaled for ease in scoring by using numbers 1, 2, 3, 4, and 5, respectively, and a question mark (?) for the "undecided" category. In the following example third-grade pupils are rated by their teacher on the following questions:[12]

Social behavior	*Rating*					
1. Does he feel secure in the group?	1	2	3	4	5	?
2. Does he show initiative?	1	2	3	4	5	?
3. Is he selfish?	1	2	3	4	5	?
4. Is he generous?	1	2	3	4	5	?
5. Does he have a favorable attitude toward the opposite sex?	1	2	3	4	5	?
6. Does he laugh *with* rather than *at* others?	1	2	3	4	5	?
7. Can he give directions to the group so they are readily accepted?	1	2	3	4	5	?

[11] Charles H. McCloy, "Character Building through Physical Education," *The Research Quarterly,* 1:41–61, October, 1930.

[12] Many of these questions were adapted from a study by Dorothy J. Dawley, Maurice E. Troyer, and John H. Shaw, "Relationship between Observed Behavior in Elementary School Physical Education and Test Responses," *The Research Quarterly,* 22:70–76, March, 1951.

Social behavior	*Rating*					
8. Does his group like him?	1	2	3	4	5	?
9. Is he ready to do his share for his team?	1	2	3	4	5	?
10. Does he encourage others on his team to work?	1	2	3	4	5	?
11. Does he insist that others follow the rules?	1	2	3	4	5	?
12. Is he courteous to other players and adults?	1	2	3	4	5	?
13. Is he just?	1	2	3	4	5	?
14. Does he take advantage of others?	1	2	3	4	5	?
15. Does he show good sportsmanship most of the time?	1	2	3	4	5	?
16. Does he accept decisions of groups and leaders?	1	2	3	4	5	?
17. Does he give and take in a friendly way?	1	2	3	4	5	?
18. Does he show sympathy to others in the group?	1	2	3	4	5	?
19. Does he show loyalty to others in the group?	1	2	3	4	5	?
20. Is he overbearing in giving directions?	1	2	3	4	5	?
21. Does he keep trying even if his team is losing?	1	2	3	4	5	?
22. Does he do things to help the other team win?	1	2	3	4	5	?
23. Does he follow the rules even though no one is watching?	1	2	3	4	5	?
24. Does he try to find ways of getting around the rules?	1	2	3	4	5	?
25. Is he kind to other players?	1	2	3	4	5	?
26. Is he fair to other players?	1	2	3	4	5	?
27. Is he honest with other players?	1	2	3	4	5	?
28. Does he take pride in his personal cleanliness?	1	2	3	4	5	?

By circling the appropriate number in each question a total social behavior score is obtained. This can be used for comparison purposes at a later date or for such purposes as parent-teacher conferences and grading.

Some pupil rating scales are precisely explained in each category of behavior. Here are two examples of behavior traits, used in a personality scale by Irwin and Staton, which illustrate the extent to which a trait can be described[13]:

Example

Personality trait	Excellent(5)	Above average(4)	Average(3)	Below average(2)	Poor(1)
D. Attitude Toward Others	Always agreeable, cooperative, and tactful	Usually agreeable and cooperative	Often agreeable and cooperative	Agreeable only when interested, uncooperative	Disagreeable, tactless, uncooperative
H. Capacity for Leadership	Very able leader	Good leader	Good leader if interested in the group activity	Poor leader	Cannot take leadership responsibility

In the construction of rating instruments character traits need to be carefully stated. Realizing this, McCloy and Hepp set out to isolate the general

[13] Leslie W. Irwin and Wesley M. Staton, *Health and Fitness Workbook*, Boston, D. C. Heath and Company, 1958, p. 153.

factors or components of which more specific traits are composed.[14] An extensive factor analysis was carried out on 142 character variables in three different studies. Fifteen factors were arrived at. The traits listed below have been found to measure the factor most satisfactorily. The italicized traits correlate quite well with the factors as indicated by the multiple correlation coefficients at the end. Thus, they should receive serious consideration when factors are being chosen to measure social efficiency.

1 "Individual-social" or individual good citizenship: *integrity, initiative,* trustworthiness, thoroughness, *conscientiousness,* poise, *ability to exert a wide influence on others, ability to work hard on one's studies,* dependability, and honesty. $R = .86$

2. "Group-social" or group good-fellowship: *ability to be a good follower, fairness, integrity,* respect for the rights of others, loyalty, perseverance, cooperation, *desire to be liked by others,* and, usually, *fondness for large social gatherings.* $R = 90$

3. Self-mastery: self-denial, moral courage, sociability, cool-headedness, *ability to recover readily from anger, striving to excel in the field of one's chief interest,* and frequently better than average athletic ability $R = 81$

4. Individual self-sufficiency: *moral courage, modesty, good sportsmanship,* resourcefulness, and perseverance $R = 75$

5. Good adjustment: *cheerfulness, persistence, active engaging in social pleasures,* lack of tendency to become angry easily or to crave admiration unduly or to give way to fits of depression $R = 51$

6 Individuality, "standing out from the crowd": *athletic ability, a good appearance,* popularity, sociability, good health, poise, and *tendency to work hard.* $R = .80$

7. Sociability: *tendency to work hard at pleasures, quickness of apprehension,* exertion of wide influence on others, originality of ideas, and ability to influence one's intimates. $R = 68$

8. Conscientiousness: *thoroughness, perseverance,* accuracy, diligence, *self-sufficiency,* good judgment, attentiveness, and ability to concentrate on one's tasks. $R = 67$

9. Buoyant individuality: *impulsive kindness, cheerfulness, quickness of apprehension,* kindness on principle, and usually a good sense of humor. $R = .66$

10. Positive action tendencies: *energy,* aggressiveness, adaptability, conviction, *resourcefulness,* initiative, *ability to discipilne others,* and usually originality of ideas $R = .90$

11. Positive attitudes and energy: *self-confidence, energy, enthusiasm,* vitality, and alertness. $R = .69$

12. Leadership: *fairness, aggressiveness, popularity,* enthusiasm, resource-

[14] Charles H. McCloy and Ferene Hepp, "General Factors or Components of Character as Related to Physical Education," *The Research Quarterly,* 28:269–278, October, 1957.

fulness, *desire to excel in the activities of one's interest, cooperation,* and *conscientiousness.* $R = .79$

13. Egotistical self-sufficiency: *eagerness for admiration, self-esteem.* $R = .56$

14. Passive group responsiveness: *cheerfulness, generosity, good-fellowship, group-mindedness, open-mindedness.* $R = .69$

15. Antisocial dominance: *readiness to become angry, desire to impose will on others, eagerness for admiration, belief in own powers, esteem of self as a whole, offensive manifestation of self-esteem.* $R = .95$

Blanchard's Behavior Rating Scale

Blanchard followed the early work of McCloy and built a rating scale consisting in twenty-four personal information items, which were listed under McCloy's nine original character traits.[15] Test reliability is .71. The intercorrelation of one trait with others in the same category is .93. Care should be exercised when using the scale not to rate too high those pupils of whom the instructor is fond. This eliminates the "halo" effect. Also it is good practice to let the pupils rate themselves occasionally. Their score can be readily compared with the score of the instructor for greater meaning (see Figure 15-1).

Cowell Social Behavior Trend Index

The purpose of this index is to discriminate and quantify the degree of social adjustment in high school youth and to provide a tentative index scale for this purpose.[16] Cowell felt that a better test of social adjustment could be made if the combined results of several competent observers were evaluated. This would occur after they responded by judging specific ways of behaving of persons in group situations. In validating the scale Cowell isolated thirty-nine descriptive behavior trends that had a high index of discriminating power for selecting socially well-adjusted and maladjusted boys. These trends and their paired opposites were evaluated by thirty teachers, resulting in the selection of twelve paired behavior trends. Further search for an "adjustment syndrome" was carried out by the Thurstone method of factor analysis and resulted in ten behavior-trend pairs being retained. Two forms of the index are available: Form A (positive traits) and Form B (negative traits). Both forms are reproduced in Figure 15-2, and the norm table appears in Table 21, Appendix A.

In using the index both forms are completed by three different teachers. The index score is the total of the ratings for all three teachers combining

[15] B. E. Blanchard, "A Behavior Frequency Rating Scale for the Measurement of Character and Personality in Physical Education Classroom Situations," *The Research Quarterly*, 7:56–59, May, 1936.

[16] Charles C. Cowell, "Validating an Index of Social Adjustment for High School Use," *The Research Quarterly*, 29:7–18, March, 1958. Forms A and B are available for $2.50 per 100 copies from the author, Purdue University, Lafayette, Ind.

Personal information	Frequency of observation						
	No opportunity to observe	Never	Seldom	Fairly often	Frequently	Extremely often	Score
Leadership:							
(1) Popular with classmates		1	2	3	4	5	
(2) Seeks responsibility in classroom		1	2	3	4	5	
(3) Shows intellectual leadership in the classroom		1	2	3	4	5	
Positive active qualities:							
(4) Quits on tasks requiring perseverance		5	4	3	2	1	
(5) Exhibits aggressiveness in his relationship with others		1	2	3	4	5	
(6) Shows initiative in assuming responsibility in unfamiliar situations		1	2	3	4	5	
(7) Is alert to new opportunities		1	2	3	4	5	
Positive mental qualities:							
(8) Shows keenness of mind		1	2	3	4	5	
(9) Volunteers ideas		1	2	3	4	5	
Self-control:							
(10) Grumbles over decisions of classmates		5	4	3	2	1	
(11) Takes a justified criticism by teacher or classmate without showing anger or pouting		1	2	3	4	5	
Cooperation:							
(12) Is loyal to his group		1	2	3	4	5	
(13) Discharges his group responsibilities well		1	2	3	4	5	
(14) Is cooperative in his attitude toward his teacher		1	2	3	4	5	
Social action standards:							
(15) Makes loud-mouthed criticism and comments		5	4	3	2	1	
(16) Respects the rights of others		1	2	3	4	5	
Ethical social qualities:							
(17) Cheats		5	4	3	2	1	
(18) Is truthful		1	2	3	4	5	
Qualities of efficiency:							
(19) Seems satisfied to "get by" with tasks assigned		5	4	3	2	1	
(20) Is dependable and trustworthy		1	2	3	4	5	
(21) Has good study habits		1	2	3	4	5	
Sociability:							
(22) Is liked by others		1	2	3	4	5	
(23) Makes a friendly approach to others in the group		1	2	3	4	5	
(24) Is friendly		1	2	3	4	5	

Fig. 15-1. The Blanchard Behavior Rating Scale

FORM A

Date:______________ Grade:______

School:______________ Age:______

Last name ______ First name ______ Describer: ______

INSTRUCTION:—Think carefully of the student's behavior in group situations and check *each behavior trend* according to its degree of descriptiveness.

Behavior trends	Descriptive of the student: Markedly (+3)	Somewhat (+2)	Only slightly (+1)	Not at all (+0)
1. Enters heartily and with enjoyment into the spirit of social intercourse				
2. Frank, talkative and sociable, does not stand on ceremony				
3. Self-confident and self-reliant, tends to take success for granted, strong initiative, prefers to lead				
4. Quick and decisive in movement, pronounced or excessive energy output				
5. Prefers group activities, work or play, not easily satisfied with individual projects				
6. Adaptable to new situations, makes adjustments readily, welcomes change				
7. Is self-composed, seldom shows signs of embarrassment				
8. Tends to elation of spirits, seldom gloomy or moody				
9. Seeks a broad range of friendships, not selective or exclusive in games and the like				
10. Hearty and cordial, even to strangers, forms acquaintanceships very easily				

FORM B

Date:______________ Grade:______

School:______________ Age:______

Last name ______ First name ______ Describer: ______

INSTRUCTION:—Think carefully of the student's behavior in group situations and check *each behavior trend* according to its degree of descriptiveness.

Behavior trends	Descriptive of the student: Markedly (−3)	Somewhat (−2)	Only slightly (−1)	Not at all (−1)
1. Somewhat prudish, awkward, easily embarrassed in his social contacts				
2. Secretive, seclusive, not inclined to talk unless spoken to				
3. Lacking in self-confidence and initiative, a follower				
4. Slow in movement, deliberative or perhaps indecisive. Energy output moderate or deficient				
5. Prefers to work and play alone, trends to avoid group activities				
6. Shrinks from making new adjustments, prefers the habitual to the stress of reorganization required by the new				
7. Is self-conscious, easily embarrassed, timid or "bashful"				
8. Tends to depression, frequently gloomy or moody				
9. Shows preference for a narrow range of intimate friends and tends to exclude others from his association				
10. Reserved and distant except to intimate friends, does not form acquaintanceships readily				

Fig. 15-2. Cowell Social Behavior Trend Index (Forms A and B).

the two forms. Raw scores are then transposed to percentile values. A high positive percentile indicates high social adjustment, and a low negative score indicates just the opposite.

New York State Social Efficiency Scale

This empirically constructed scale was designed by the New York State Education Department and consists of two parts: social efficiency and cultural appreciation.[17] As a measure of individual progress in social adjustment it is administered twice about a month apart, for comparison purposes. Nine categories dealing with initiative, perseverance, courage, self-control, self-reliance, cooperation, fair play, sympathy, and sincerity are appraised and graded on a point system: "usually" (3), "often" (2), "occasionally" (1), and "seldom" (0).

Carr Attitude Scale

Employing descriptive statements on social, personal, and activity attitudes, Carr designed an eighty-four statement social attitude scale primarily for high school girls.[18] It was validated on the ratings given by ninety-six physical education judges. In the scale there are sixteen most desirable statements, twenty-one desirable statements, twenty-eight undesirable statements, and nineteen least desirable statements. Scoring is accomplished by subtracting the percentage of undesirable attitudes from the percentage of desirable attitudes. The complete scale appears in the original source.

Winnetka Scale

Constructed to rate the social and emotional aspects of the personality of school children from nursery school through grade six, the Winnetka Scale appraises personal cooperation, social consciousness, emotional security, leadership, and responsibility.[19] It has been used to diagnose personality difficulties at an early age. This is a situation scale, where thirteen situations are set up, with a number of choices of action listed under each. Validity and reliability are satisfactory.

Wear Attitude Scale

This scale was previously mentioned in Chapter 12 in connection with attitude measurement in physical education. The two equivalent forms, each with thirty statements, were designed by Wear to sample attitudes toward

[17] The University of the State of New York, *Secondary Physical Education Syllabus*, Albany, N.Y., State Department of Education, Division of Health and Physical Education, Bulletin no. 1062.

[18] Martha G. Carr, "The Relationship between Success in Physical Education and Selected Attitude Expressed by High School Freshmen Girls," *The Research Quarterly*, 16:176–191, October, 1945.

[19] Dorothy Van Alstyne, "A New Scale for Rating School Behavior and Attitudes in the Elementary School," *Journal of Educational Psychology*, 27:677–704, December, 1936.

physical education in general, but many of the questions relate to social behavior.[20] Both forms of the scale appear in the cited reference.

SITUATION RESPONSE TESTS

As already indicated there are weaknesses in the type of statement that must be answered simply "yes" or "no." There are also weaknesses in rating scales where a category of behavior is judged by the instructor. To overcome this the situation response has been used, in which the subject usually responds to a well-described situation.

In *The Sportsmanship Preference Record* the subject responds to twenty situations commonly encountered in the junior high school physical education classes.[21] Two examples from this test are as follows:

Examples

It is against the rules to pitch side arm, but you know a way to pitch side arm that is hard to detect. If you pitch side arm you know you could win this game and your team would be school champs. What would you do?

(*a*) Pitch side arm until the umpire catches you.
(*b*) Pitch side arm just when the good hitters are up to bat.
(*c*) Don't pitch side arm at all.
(*d*) Don't pitch side arm unless you start to lose.

You are pitcher on your softball team. The team captain knows this but he wants to be pitcher himself, so he has you play first base. You want to pitch and you don't like the first base.

(*a*) Refuse to play unless you get to pitch.
(*b*) Do a poor job of playing first base so that you will get changed to something else.
(*c*) Get the other members of your team to tell the team captain you are the best pitcher.
(*d*) Play first base as best you can.

In this test there was found to be a significant difference to the 5 per cent level of confidence between sportsmanship attitudes of sixth-, seventh-, and eighth-grade boys. On a retest after six months a reliability coefficient of .80 was found between the original and final tests. A correlation coefficient of .53 was found between the physical education teachers' evaluation of the boys' sportsmanship attitudes and the test results. Copies of the test can be obtained from the author at Monroe Junior High School, Inglewood, California.

SOCIAL ADJUSTMENT INVENTORIES

By requesting a pupil to take a look at his problems and interests it is possible to arrive at an inventory of strong and weak characteristics. Social ad-

[20] C. L. Wear, "Construction of Equivalent Forms of an Attitude Scale," *The Research Quarterly*, 26:113–119, March, 1955.
[21] Robert A. McAfee, *op. cit.*, p. 120.

justment inventories vary in construction, but most attempt to point out the maladjusted person. As a rule, they are better screening devices than they are diagnostic instruments. It is in this connection that they have their greatest value to physical educators. However, the most practical use of these test results is gained when the guidance specialist and the physical instructor sit down together to interpret individual cases.

Science Research Associates Inventories (SRA)

Two useful adjustment inventories for boys and girls have been prepared by Science Research Associates.[22] The first is the SRA Junior Inventory for elementary grades four to eight. The second is the SRA Youth Inventory for secondary grades seven to twelve. Both inventories relate to pupil problems that the pupils say bother them the most. The inventory is used by having the pupil survey a long list of possible problems (223 for the junior inventory and 298 for the youth inventory) and check off personal problems. The number of checked items represents the score in each area.

Bell Adjustment Inventory

Two forms of this inventory are available: one for high school and college students, the other for adults.[23] Five separate scales of personal and social adjustment are provided: home, health, social, emotional, and occupational. The occupational scale appears only in the adult form. Reliability for the total score is .94. Only test items that survived analysis for discriminatory power were kept in the battery. Further test validation was accomplished by specialists who selected "very well" and "very poorly" adjusted groups for inventory differentiation. A point in favor of the Bell Adjustment Inventory is that each adjustment item has a descriptive designation that can be rated to scale (excellent, good, average, unsatisfactory, very unsatisfactory). It has been used in a number of institutions by physical education and guidance personnel in an effort to gain insight into the immediate adjustment problems of physical education students.

Bernreuter Personality Inventory

This inventory is designed to measure personality traits having to do with neurotic tendencies, self-sufficiency, introversion-extroversion and dominance-submission.[24] It consists of 125 "yes," "no," or "?" questions applicable to secondary school and college pupils as well as to adults. Over the years some improvements have been made on the test. Two additional scales, "self-confidence" and "sociability," were added by Flannigan. It may not be necessary to use the whole test, because the traits on the neurotic-tendency

[22] Science Research Associates, Inc., 57 Grand Ave., Chicago, Ill.
[23] Hugh M. Bell, *The Theory and Practice of Student Counseling*, Stanford, Calif., Stanford University Press, 1935.
[24] Robert G. Bernreuter, "The Theory and Construction of the Personality Inventory," *Journal of Social Psychology*, 4:387–393, November, 1933.

scale correlate highly with the traits on the introversion-extroversion scale. In fact, Bernreuter points out that the most valuable part of his inventory is the neurotic-tendency scale.

Minnesota Multiphasic Personality Inventory

There are a great many tests of personality. In 1937, Ruth Strang found that there have been as many as 4,000 personality tests. Obviously, not all these are worthwhile. One of the better ones is the Minnesota Multiphasic Personality Inventory, developed by Hathaway and McKinley.[25] Here the high school or college student is asked to read each of 566 short statements and decide whether each one is "true as applied to you" or "false as applied to you." An analysis of the answers according to the numbered statements permits a multiphasic diagnosis of strengths and weaknesses in personal adjustment. Profile scale sheets are completed for each sex.

California Psychological Inventory

This inventory consists of eighteen standard scales relating specifically to dominance, capacity for status, sociability, social pressure, self-acceptance, well-being, responsibility, socialization, self-control, tolerance, good impression, communality, achievement via conformity, achievement via independence, intellectual efficiency, psychological-mindedness, flexibility, and femininity.[26] It is a good example of a diagnostic instrument to be administered and interpreted by the psychological or guidance specialist. The relationship of motor ability to the eighteen separate scores on this inventory was investigated by Keogh.[27] In this study, the CPI scores of 167 college students did not correlate significantly with either motor ability or athletic participation. This would suggest that one should be careful in generalizing about personality characteristics and athletic participation. It might also suggest a closer look at the nature of the sample investigated. Moreover, many variables have a bearing on personality, some of which are not highly related to athletic or motor ability.

Mooney Problem Check List

An easy listing of problems to administer and score was developed by Mooney for college students showing signs of social maladjustment.[28] It is a grouping of 330 "troublesome problems which often face students in college—problems of health, money, social life, relations with people, religion,

[25] Starke R. Hathaway and J. Charnley McKinley, *Minnesota Multiphasic Personality Inventory*, New York, The Psychological Corporation, 1943; revised, 1948.

[26] Harrison G. Gough, *Manual for the California Psychological Inventory*, Palo Alto, Calif., Consulting Psychologists Press, 1957.

[27] Jack Keogh, "Relationship of Motor Ability and Athletic Participation in Certain Standardized Personality Measures," *The Research Quarterly*, 30:438–445, December, 1959.

[28] Ross L. Mooney, *Problem Check List*, Bureau of Educational Research, The Ohio State University, Columbus, Ohio, 1941.

studying, selecting courses, and the like." The student is asked to select and underline the particular problems of concern to him. He then goes over the problems again indicating which are of most concern. After completing these steps, he answers five summarizing questions of the essay type, which inquire into his reactions to the check list and what is to be accomplished by using it.

The instructor using the results of the Mooney check list can see at a glance the trend of thirty questions in each of eleven areas of emphasis. The author has found this inventory of problems most helpful in working with adapted physical education students.

Washburne Social Adjustment Inventory

In this inventory Washburne has met the need for a test to discover children with differences in adjustment without the complication of differences in intelligence.[29] Over a ten-year experimental period Washburne tried out numerous questionnaires on groups of pupils, matched for intelligence, age, and sex, but contrasted for social behavior. The final form of the inventory was validated as follows: bright, adolescent, third- or fourth-offense prisoners, called socially maladjusted on the basis of their records, were paired and contrasted with unselected high school students; high school boys and girls judged exceptionally maladjusted by three competent judges were paired and contrasted with boys and girls judged exceptionally well adjusted. From over 5,000 testings, 400 pairs were selected. When the social adjustment scores of these four groups were plotted, the result was four overlapping normal curves, arranged to indicate high test validity. The coefficient of validity (biserial r) is .90. Reliability is .92, and the correlation with intelligence is .07.

The Washburne inventory is a group test for all ages above the eighth grade. It consists of 123 items arranged in questionnaire form, which can be answered in approximately thirty to fifty minutes by a simple "yes" or "no" response. The test is self-explanatory and contains eight subtests, with the following names, which were chosen as being the nearest approach to a one-word description of the common factor in a group of questions: (*t*) truthfulness; (*h*) happiness; (*a*) alienation; (*s*) sympathy; (*p*) purpose; (*i*) impulse-judgment; (*c*) control; (*w*) wishes. With the exception of test *w*, which is the essay type and of importance chiefly for purposes of interviewing, the subtests are completely objective and can be scored by a simplified stencil key or by machine.

The inventory is suitable for use in health and physical education, where pupils can be selected as in need of befriending and special counsel. It has also been used by playground supervisors and by camp counselors to appraise

[29] John N. Washburne, *Social Adjustment Inventory*, Yonkers, N. Y., World Book Company, 1940; see also John N. Washburne, "A Test for Social Adjustment," *Journal of Applied Psychology*, 19:201–214, April, 1935.

changes in social adjustment through precamp and postcamp measurement. Norms and a profile chart are provided so that a profile of a pupil can be drawn to show departures from the average. Also, the first page of the test booklet provides space for information questions that should be of help to the adviser; these relate to student interests, socioeconomic status, and home conditions in general.

The close relationship between social adjustment and physical education skills and interests was brought out at Syracuse University, when the major students in three successive junior classes rated above the seventy-fifth percentile on Washburne's national norms.

SOCIOMETRIC TECHNIQUES

It is not the intent of this chapter to make sociometrists or measurement psychologists out of health and physical education teachers. The emphasis is solely on familiarity with some of the available tools and techniques and the team approach in seeking answers. In this connection sociometry holds considerable promise.

Sociometry is a relatively new technique of evaluation, pioneered by Moreno.[30] His work, together with that of Jennings, has been adapted for successful use in industry, business, education, and recreation.[31]

Sociometry is a quantitative method of studying the organization of groups. It presents graphically the entire structure of relations present at a given moment among the members of a particular group. It explores the natural friendship patterns that exist among members. Using this technique. the instructor may gain increased knowledge of the interpersonal relationships with a class of pupils and discover which youngsters are accepted by their own age group and which are not. Very often the latter category need special help in making wholesome social adjustments. These pupils, often on the fringe of the group and in an unwanted role as "isolates," need help both in the classroom and on the playground. Todd reinforces this viewpoint by praising the use of sociometry in physical education.[32] In fact, at the high school level Todd carried out a sociometric study and obtained objective data showing greatly increased acquaintanceship and a significant decrease in the number of unpopular and unwanted girls during a one-semester experimental period.[33] Moreover, anonymous questionnaires revealed that the girls found the sociometrically selected squads more enjoyable and efficient than any other grouping method they had ever experienced.

[30] J. L. Moreno, *Who Shall Survive?* New York, Beacon House, Inc., 1934.

[31] Helen Hall Jennings, *Sociometry in Group Relations,* American Council on Education, Washington, 1950.

[32] Frances Todd, "Sociometry in Physical Education," *Journal of Health, Physical Education and Recreation,* 24:23–25, May, 1953.

[33] Frances Todd, "Democratic Methodology in Physical Education," unpublished doctoral dissertation, Stanford University, Stanford, Calif., 1951.

A number of studies support the validity and reliability of sociometric measures. Moreno states that the tests are self-validating if rapport is good. Cook summarized several studies of reliability that indicated coefficients between .90 and .96.[34] Prentice extolls the use of these measures in physical education and points to their validity.[35] Bonney and Fessenden reviewed twenty studies concerning reliability and found that the median rank-order coefficient was .90 when the time interval of the retest was no more than one week.[36] Incidentally, the later the retest was given, the lower the coefficients of reliability. This is to be expected, for as children work together, they tend to make new acquaintances and change opinions toward others.

An interesting way of checking the validity of sociometric measures is to compare them with the relatively more established scales and inventories of social adjustment and such other data as can be obtained from observation records, pupil ratings of each other, and teacher ratings of pupils. In this respect Bonney and Fessenden found that, when the two types of tests were properly compared, the correlation was not great, but when those pupils high in the upper quartile in choice status are compared with those in the lower quartile, the findings are generally consistent in showing the most frequently chosen persons to be superior to the infrequently chosen pupils in some recognized types of behavior adjustments.[37]

Kinds of Sociometric Techniques

Sociometry can be employed at any grade level. Bonney and Burleson stress its use in classes of physical education at the junior and senior high school because it appears here most readily to foster good student interpersonal relations.[38] Actually, the problem is not so much where to use sociometry as the nature of the technique employed. Todd refers to two techniques:

1. **The Acquaintance Volume Test.** This is an objective means of determining how well students get acquainted in a semester or unit. At the first class meeting the pupils are asked simply to write the first and last names of those they know in the group. At the end of the term or unit this is repeated, and by easy arithmetic it is apparent how many new friends each person has made.

2. **The Functional Choice Test.** Here it is determined who *wants* to be with whom and not just who *is* with whom. Group members are given a

[34] Lloyd A. Cook, "Manual for Obtaining, Analyzing and Diagramming Sociometric Data," Detroit College Study in Intergroup Relations, Detroit, Mich., 1950.

[35] M. Erma Prentice, "An Experience in Group Dynamics," *Journal of Health, Physical Education and Recreation,* 24:24–26, October, 1953.

[36] Merl E. Bonney and Seth A. Fessenden, *Manual: Bonney-Fessenden Sociograph,* Los Angeles, Calif., California Test Bureau, 1955.

[37] *Ibid.,* p. 47.

[38] Warren C. Bonney and Reba M. Burleson, "Socializing Techniques," *Journal of Health, Physical Education and Recreation,* 25: 40–44, March, 1954.

chance to choose or reject others. This works well when each sociometric choice makes sense to the person making the choice; that is, it must involve something he wants to do, and it should carry with it the promise that his wishes will be taken into account in making changes.

Application of Sociometry to Health and Physical Education

By using functional choice tests in health and physical education classes it is possible to find out early which students need the immediate attention of the instructor and how they should be grouped for active participation in the program.

When the class has been together long enough to know each other fairly well, the teacher may casually ask the pupils to write their first three preferences for squadmates for an activity on a 3- by 5-inch file card. The students should be promised that their choices will be kept confidential and that the teacher will guarantee to place each pupil on a squad with at least one of his chosen friends. Then it should be suggested that, if there is anyone with whom the student would prefer *not* to play, the name or names of that person be written at the bottom of the card. With the cards collected and this information on hand, there are several possible procedures for tabulating and using the data:

1. Study the cards and set up the squads directly, or
2. Build a matrix chart
3. Construct a sociogram
4. Evaluate the individual scores

The Matrix Chart

To construct a simple graphical picture or matrix chart, the names and numbers are listed down the chart as well as across it. The average, the popular, and the unpopular pupils can then be seen at a glance (see Figure 15-3).[39] An acceptance score for each pupil is included in the tabulation chart to make the choices more meaningful. This is a weighted score figured with 3 points for the first choice, 2 for the second, and 1 for the third choice. Ann, for example, had three first choices and two second choices for a total of 13 acceptance points. An examination of Figure 15-3 reveals the number of times each pupil chose, whom he chose, the acceptance scores for each, and the total number of choices used. By itself the chart is useful.

The Sociogram

Another procedure is to use the information from the tabulation chart to build a *sociogram*. A sociogram shows the highly differentiated all-per-

[39] Reproduced by permission of Patricia Whitaker Hale, Santa Barbara College, Goleta, Calif. See "Proposed Method for Analyzing Sociometric Data," *The Research Quarterly*, 27:152–161, May, 1956.

vading network. It shows the "stars," the unchosen, and a few mutual pairs. As groups become better acquainted, more pairs are formed. The sociogram also shows the "isolates," persons who choose no one, who are not chosen, or who have no reciprocated choices. The fewer isolates in the group, the better; the higher the total number of reciprocated choices, the better. As groups mature, the number of overchosen pupils decreases and the choices are shared by all.

Choice	Ann	Betty	Carol	Dorothy	Enid	Frances	Greta	Helen	Irene	Janis	Kathie	Lois	Margie	Norma	Pat	Ruth	Choices used
Ann	x									3	1					2	3
Betty		x						2						1			2
Carol	2		x													1	2
Dorothy				x		1				3						2	3
Enid			1		x										2		2
Frances						x			3	2						1	3
Greta							x										0
Helen	1							x						2			2
Irene						1			x	3						2	3
Janis						2				x			3			1	3
Kathie	1										x	2					2
Lois	1				3						2	x					3
Margie				2						1			x			3	3
Norma		1												x		2	2
Pat			1		3										x	2	3
Ruth	2					3				1						x	3
1st	3	1	2	0	0	2	0	0	0	2	1	0	0	1	0	3	
2d	2	0	0	1	0	1	0	1	0	1	1	1	0	1	1	5	
3d	0	0	0	0	2	1	0	0	1	3	0	0	1	0	0	1	
Total	5	1	2	1	2	4	0	1	1	6	2	1	1	2	1	9	
Acceptance score	13	3	6	2	2	9	0	2	1	11	5	2	1	5	2	20	

Fig. 15-3. A Matrix Chart for Tabulating Sociometric Data.

The data from Figure 15-3 has been made a part of the sociogram in Figure 15-4. An examination of this reveals the number of reciprocated choices as well as the number of pupils receiving reciprocated choices.

A sociogram identifies cliques, pairs, threesomes, gangs, and the unwanted child. In drawing it, the popular pupils are placed in the center of the diagram and the lonely and rejected ones on the periphery. The lines of attraction and reciprocated and unreciprocated choices are then drawn in. Rejections can be charted in colored pencil. In Figure 15-4, Ruth and Ann both qualify as overchosen, and Greta, Dorothy, Carol, and Helen need to be brought into the group.

Another step that Hale has employed is to set up a rank-listing chart after a second testing and the making of a second tabulation chart.[40] The rank-listing chart reveals the range of scores on both tests, the stars and the

[40] Patricia Whitaker Hale, *op. cit.*, p. 158.

isolates, and the gains and losses made by each individual, and it provides a clear picture of the relative status of each pupil in the group. If progress in socialization is being made in the class, the most chosen pupil on the first test should have lost some choices and the least chosen should have gained some.

To structure the sociogram results so they are more helpful to the teacher, Cook has set up an *individual status index,* which is a simple formula assigning to each pupil a numerical symbol of his popularity with others.[41] By retesting during the term, it is possible to show changes in pupil status. To his tabulation chart Cook adds an *R* for each pupil rejected by anyone.

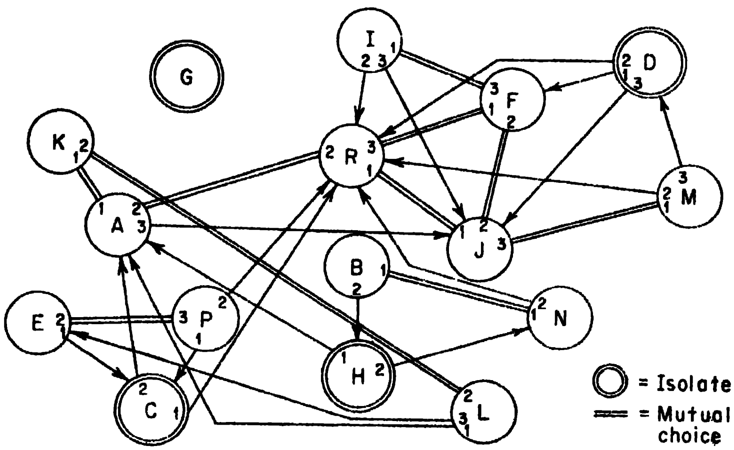


Fig. 15-4. The Sociogram. (Courtesy Patricia Whitaker Hale.)

Not being chosen is one thing, but being purposely rejected is far worse and places the pupil farther toward the periphery. The following formula is used:

$$\text{Individual status index} = \frac{\text{total choices} - \text{total rejections}}{\text{number in group} - 1}$$

For example, in the previous illustration, in a group of sixteen pupils, Ruth received nine choices and no rejections; her individual status index is .8. Carol, on the other hand, with a total of two choices and no rejections would have an ISI of .13.

Like many other measures, the sociogram needs qualified handling. Pupils should not be told that they are taking a sociometric test. This may evoke the wrong connotation. Rather they should be asked to help the teacher group them for an activity. Once the choices have been made, a grouping plan is recommended. For instance, the unchosen or seldom chosen pupil should be placed in a group with his first choice. Likewise, if a pupil is chosen by others but chooses none of those who choose him, give him his

[41] Lloyd A. Cook, *op. cit.*, p. 24.

first choice. Regardless of status each pupil should be placed with at least one of his choices. Rejections, of course, should not be grouped together.

Cowell Personal Distance Ballot

Another form of sociometric device is the personal distance scale developed by Cowell to objectify teacher judgments of social behavior.[42] Here the pupil is asked to consider each member of his group in terms of *how near* he would like to have him to his own family. There are seven choices, ranging in degree as follows:

I would be willing to accept him:

Into my family as a brother	As a very close "pal" or "chum"	As a member of my "gang" or club	On my street as a "next door neighbor"	Into my class at school	Into my school	Into my city
1	2	3	4	5	6	7

In such an arrangement the personal distance score is calculated by adding the total weighted scores given the pupil by members of the group and dividing by the total number of respondents. The low score is most desirable. Percentile scale scores for boys are included in the cited reference. Cowell, working with Trapp, at Purdue University, with the football squad, reports test reliability coefficients of .91, .88, and .93. This sociometric device has earned a place along with the sociogram as a means of appraising at least one area of social efficiency, namely, social status.

The Human Factor

The success of the sociometric technique of grouping and the progress that individual pupils make is not automatic. In the end it is related to the understanding and keen observation of the instructor. The human factor in measurement is still the greatest influence. Structured characteristics, formulas, and statistics are only means to an end.

Another human factor suggests a word of caution at this point. Not all persons conform equally to the group pattern. There will be persons who prefer, and even seek, a small circle of friends, and it is conceivable that they will be socially efficient in their own way to a satisfactory degree. It would be a mistake to attempt to place every member of the class or group in some socially structured organization or classification. There must always be room for individual deviants who may someday make a contribution to society and civilization, who may be a little less outgoing, a little less sociable, and a little more contemplative as they seek the inner self. The sensitive temperament of the cerebrotonic must not be smothered by the over-

[42] Charles C. Cowell, *op. cit.*, p. 9. This instrument, with lines for forty students, is available from the author, Purdue University, Lafayette, Ind.

ambitious viscerotonic-somatotonic urge to categorize all social behavior by group behavior or some other limited objective.

SELECTED REFERENCES

Bentson, T. B., and John Summerskill: "Relation of Personal Success in Intercollegiate Athletics to Certain Aspects of Personal Adjustment," *The Research Quarterly,* 26:8–14, March, 1955.

Breck, Sabina J.: "A Sociometric Measurement of Status in Physical Education Classes," *The Research Quarterly,* 21:75–79, May, 1950.

Buros, Oscar K. (Ed.): *The Fifth Mental Measurement Yearbook,* Highland Park, N.J., The Gryphon Press, 1959.

Cowell, Charles C.: "An Abstract of a Study of Differentials in Junior High School Boys Based on the Observation of Physical Education Activities," *The Research Quarterly,* 6:129–136, December, 1935.

Fraleigh, Warren P.: "The Influence of Play upon Social and Emotional Adjustment with Implications for Physical Education," *Proceedings, College Physical Education Association,* 59:268–273, 1956.

Fulton, Ruth, and Elizabeth Prange: "Motor Learning of Highly Chosen and Unchosen Teammates and Friends," *The Research Quarterly,* 21:126–131, May, 1950.

Jones, Harold E.: "Physical Ability as a Factor in Social Adjustment in Adolescence," *Journal of Educational Research,* 40:287–301, December, 1946.

———: *Motor Performance and Growth,* Berkeley, Calif., University of California Press, 1949, chap. 8.

McCraw, L. W., and J. W. Tolbert: "Sociometric Status and Athletic Ability of Junior High School Boys," *The Research Quarterly,* 24:72–76, March, 1953.

Moreno, J. L.: *Who Shall Survive?* 3d ed., New York, Beacon House, Inc., 1953.

Prentice, M. Erma: "An Experience in Group Dynamics," *Journal of Health, Physical Education and Recreation,* 24:24–26, October, 1953.

Todd, Frances: "Sociometry in Physical Education," *Journal of Health, Physical Education and Recreation,* 24:23–24, May, 1953.

Watson, Goodwin: "Personality Growth through Athletics," *Journal of Health and Physical Education,* 9:408–410, September, 1938.

CHAPTER 16

Use of Tests in Meeting Individual Needs

No citizen has a right to be an amateur in the matter of physical training . . . what a disgrace it is for a man to grow old without ever seeing the beauty and strength of which his body is capable.

Socrates

The task of education is to differentiate between the fit and the partially fit in all areas of the educational program. To do otherwise is to ignore the range of individual capacities of intellect, general ability, social adjustment, and organic health. It is the act of differentiation that calls for exact measurement and evaluation. Moreover, such appraisal, to be worthwhile, must be a continuous process of relating individual and program status to educational goals.

Redirecting Physical Education

In the ninety years before World War II the building of sound organic health or physical fitness was the major function of health education and physical education. Today this is an important objective, but it is frequently subordinate to such aims as social efficiency, recreational competency, and athletic skill. In many health classes, mental health draws more attention than any other phase of health education. In some physical education circles the word "physical" is almost distasteful because of the announced fear that someone will put too much emphasis on it. The truth of the matter is that the unique function of health and physical education to deal with the typically healthy and the atypical alike is being overlooked and even relegated to the past. The basic need nationally is to motivate a return to primary consideration of individual differences in health.

The fundamental problem in raising the physical standards, and the corresponding capacity for total human activity, is to improve the fitness of those boys and girls who are subpar in basic strength and stamina, as well as those suffering from malnourishment, chronic fatigue, physical defects, organic strains, and abnormalities of hearing, vision, and body mechanics. Effective treatment and physical education programs for those

below standard should be fundamentally different from that for typical pupils or even those who are athletically inclined.

A redirection of emphasis is needed now more than ever before. The remedial, adapted, or subpar program should be designed to identify children at all ages who are below acceptable levels in physical education through increased use of appropriate tests, which will help discover the existence and causes of low physical fitness and social maladjustment. This is to be done by conducting individual case studies and carrying out appropriate follow-up. Depending on individual pupil needs, follow-up services may include developmental and conditioning exercises, advice on health habits, recreational therapy, rest, and relaxation, and referral to other specialists as the need arises.

To Clarke, the basic fitness problem exists with boys and girls who are deficient in basic strength and endurance elements and yet are organically and structurally sound and well nourished.[1] Moreover, it is this group which is generally neglected today and for which provisions have not been specifically made in most departments of physical education. The redirection of emphasis in physical education, which will be most appreciated by parents and pupils themselves, is to meet individual requirements of boys and girls who are handicapped in some respect, who have functional defects or deficiencies amenable to improvement through exercise, or who possess other inadequacies that interfere with their successful participation in the diversified and vigorous activities of the regular physical education curriculum. Clarke suggests seven basic functions of this program[2]:

1. General development and conditioning activities for individuals free of handicaps but of low physical fitness status.
2. Body mechanics training for individuals with non-pathological conditions.
3. Adaptation of physical education and recreational activities for the handicapped.
4. Psychological and social adjustments of "normal" individuals with atypical tendencies.
5. Relaxation activities for individuals suffering from chronic fatigue and neuro-muscular hypertension.
6. Counseling, guidance, and assistance with physical fitness, personal adjustment, and social problems.
7. Selected procedures for improving the motor techniques of unskilled individuals.

No such remedial effort can be properly initiated or scientifically carried through to completion short of testing and examining practices. In addition to tests of physical and social status, continued appraisal is facilitated through such devices as the interview, questionnaire, check list, anecdotal record, and case study.

[1] H. Harrison Clarke, *Physical Fitness Newsletter,* 5:5, February, 1959.
[2] *Ibid.,* p. 5.

Gifted pupils frequently need a special physical education. Failure to develop adequate strength, muscular endurance, and body coordination limits genius. Poor health and lowered physical capacity reduce ability to perform as a genius or as a moron. We approach our potential mental capacity only when we are capable of putting our thoughts into action. This can be illustrated in a number of ways. For example, in a first-grade study of Caucasian pupils drawn from all socioeconomic levels, it was found that failing students tended to be more immature on a battery of anthropometric indices than successful students.[3] The difference did not depend entirely on age or IQ, as failing students matched for age and IQ with the top students were also found to be immature. The implication is that anthropometric measures may help identify immature pupils who have trouble keeping up with the group, both physically and mentally.

Handling Exceptional Children

It is not the intent of this chapter to investigate the varieties of exceptional children or to review all the tests and measures that might well be used in this program. Rather it is the purpose to show how measurement can be applied specifically to the adapted health and physical education program.

The extent to which exceptional children frequent such programs is evidenced by the growing numbers of pupils who are (1) physically handicapped, (2) mentally retarded, (3) gifted, (4) behavior problems, and (5) substrength. A cross section of adapted physical education classes reveals a wide variety of complaints including postural defects, bone and joint disorders, malnutrition, rheumatic heart, cerebral palsy, poliomyelitis, epilepsy, visual and auditory handicaps, anemia, asthma, inguinal hernia, postoperative cases, diabetes, and emotional instability.

Classification and Assignment to Adapted Physical Education

Medical diagnosis serves as a basis for the adapted program This should be coupled with a plan of cooperation between the school physician and the physical education instructor. These two specialists working together can accomplish much individual pupil rehabilitation. The physician should assume the leading role.

Pupils may be assigned to adapted classes on the basis of medical examinations, screening tests, physical capacity and achievement tests, parent recommendations, observation and inspection data from teachers, and data from local clinics and health agencies.

Yearly more students with physical deviations join adapted classes. In one large Midwestern university, for example, there was an increase of over 38 per cent in a three-year period and an increase of 56 per cent in the number

[3] Maria D. Simon, "Body Configuration and School Readiness," *Child Development*, 30:72–75, December, 1959.

of handicapped students with disabilities.[4] The increase is noted in secondary education as well, with more remedial physical education specialists being employed each year. In one survey of 106 cities of 100,000 or more population, ninety-eight school departments provided special programs for handicapped children.[5] Orthopedic handicaps are the most broadly provided for.

Evaluation Practices in the Adapted Program

The list of evaluation practices that can be employed in an adapted physical education program is long indeed. Tests of capacity, skill, achievement, body mechanics (including the feet), social efficiency, and body build all have a place. A few of the more useful devices and techniques will be mentioned.

Performance Testing. Physical fitness and motor ability tests reveal who is in need of special attention. Substrength or subperformance may be only a symptom of something more serious, such as organic disease, poor feet, or malnutrition. In this respect PFI tests have proven most reliable over the years in charting individual progress.

A recommended procedure for meeting individual differences, which was worked out at Boston University and Syracuse University and applied in a number of secondary schools, is as follows[6]:

1. Medical examination by university physician. Students possessing handicapping defects or defects of such a nature that vigorous exercise would be harmful to them are reported with recommendations to the Department of Physical Education where appropriate individual programs of activity are designed.
2. Administering of the Physical Fitness Index Test to determine status.
3. The selection of those individuals low in physical fitness requiring special study.
4. A personal interview with each student in an effort to ascertain the cause or causes of low fitness.
5. Assign individuals to physical education classes according to the basis of need.
6. A study of medical history, personal health habits, social and psychological adjustment, and scholastic records.
7. Initiation of appropriate remedial and corrective procedures.
8. The periodic re-testing of students to determine progress, and to adapt programs to the findings.

The results of this technique can best be described in terms of actual outcomes. The following case study is presented:

Robert G., a senior in the College of Liberal Arts, was doing good college work. At the time of college entrance his physical fitness index score was 82. Initial study into the underlying causes of this low physical fitness revealed the presence

[4] Agnes M. Hooley, "We Can Serve the Students with Disabilities," *Journal of Health, Physical Education and Recreation*, 30:45–46, March, 1959.

[5] Helen M. Wallace and Helen M. Starr, "School Services for Handicapped Children," *American Journal of Public Health*, 50:173–180, February, 1960.

[6] Carl E. Willgoose, "Meeting Individual Physical Fitness Needs," *School and Society*, 67:446–447, June 12, 1948.

of several limiting factors, none of which, however, seemed severe enough to bring about such a low physical fitness rating.

The student was placed in a modified physical education program of mild activities, and progress was observed very closely during his first semester of college life. When no improvement in fitness was found at the end of his first semester, he was strongly advised to see his family doctor for a thorough medical examination. The physician, after careful study and numerous tests, reported that Robert was suffering from a serious thyroid condition and advised immediate surgery. The operation was successful and the condition corrected. The medical doctor reported that the nature of the defect was so serious that, had it not been corrected when it was, it might have resulted in death. It is an excellent example of the importance of studying carefully those students indicating undesirable deviation from normal standards.

The Case Study Approach. There is no more fruitful way to understand individual pupil differences than the case study technique. Here the teacher is the prime mover in following up basic health abnormalities. The objective is to get at the cause through careful interviews with teachers, parents, and the subject and through examination of medical and scholastic records. When the academic record is viewed in the light of height-weight records, anecdotal records, the health history, recent illnesses, functional impairments, and personal health habits, the combined information can be most enlightening. For ease in assembling this kind of information it is recommended that the Clarke Health-Habit Questionnaire and Case Study Form be used. For a review of the questionnaire, see Chapter 5.

Frequently the case history is kept near at hand for immediate use by the physical education instructor. It will usually include the results of the medical examination and the special recommendations of the supervising physician, the results of special consultations, treatments, referrals, and student autobiographies and appraisals. In this connection Kent State University has its students in adapted physical education write two papers, which are included in the file.[7] The first paper consists of a brief biography, a discussion of the causes and effects of his present condition, the possible adjustments that it will be necessary for him to make, his life ambition, objectives he hopes to attain, and what he expects from the program. The second paper is written at the end of the quarter. It includes an evaluation of the course, the student's achievements, skills that he has learned, possible change of outlook, other accomplishments, and criticisms of his participation in the adapted physical education section. Needless to say, this provides the instructor with another means of appraising success in the adapted program.

Individual Conferences. As an evaluation technique the individual conference or interview has considerable merit, provided the correct atmosphere exists. Proper rapport between pupil and teacher is needed if questioning and counseling relative to habits of living is to be effective. In recent years,

[7] J. A. Fischer, "Helping to Solve the Social and Psychological Adjustment Problems of the Handicapped" *Journal of Health, Physical Education and Recreation,* 31:34–36, February, 1960.

the *structured interview* has been employed to pinpoint questions more clearly and make the conference more objective.

Health Councils. Getting at the basis of human behavior is no longer something one can do alone. It requires a team approach, which frequently follows health appraisals. The combined thoughts of several persons are brought to bear on the health needs of a particular child, often under the chairmanship of the school health supervisor or health coordinator. Groups such as these permit the school to give more than "lip service" to meeting individual needs. A council such as this often finds among its members the classroom teacher, nurse, physician, and physical education instructor.

The health council draws on all sources of information, much of which has been gathered by the nurse or physical education teacher. Information is weighed in the light of known circumstances: home life, community activities, etc. Fitness is interpreted to some extent in terms of individual constitutional endowment. Endomorphic extremes, for example, are not expected to score as well on tests of strength, endurance, and motor ability as the mesomorphs. It is conceivable, therefore, that a low mesomorphic performance score is indicative of a more serious inadequacy in physical capacity than the low endomorphic score. In any case, the health council is the place to evaluate the evidence and suggest help.

Health Inventories and Check Lists. As already pointed out in Chapters 5 and 15, a number of inventories and check lists exist to help the health and physical education teacher sample the status of boys and girls. Health attitude scales, such as those by Byrd or Mayshark, and check lists that reflect feelings toward physical education are sometimes very useful in affording "fill-in" details. Inventories, such as those by Neher, Johns and Juhnke, Leonard and Horton, and Mooney, may be similarly used. The point of emphasis, of course, is that all available means should be used to improve the health of the exceptional child.

Record Keeping. Everything from the medical findings to the written opinion of a parent is included in records. As the schools become more highly organized and complex because of increased school enrollments, it is necessary to maintain accurate records. The magnitude of this statement is vividly brought out by the fact that there is a 30,000 pupil turnover every year in Los Angeles alone.

A comprehensive folder, prepared for each pupil in adapted physical education, should contain much of the information already referred to and be kept for confidential staff use. Such a folder, in addition to its chief purpose in physical education, can be used to obtain additional pertinent data for the cumulative health record. Although there are no standardized sets of records, there are a number of well-organized record folders in many of the larger school systems that have been fashioned by trial and error. The record for the corrective physical education section in the Los Angeles City School Districts is a regular manila folder, which opens up to two full sides of information, as follows:

Inside left:

Pictures: full length, "first," "second," and "third," with dates taken.

Special individual exercises: ample room for instructor to be, specific in details.

Corrective teacher's notes: room for comments.

Inside right:

Screening (first semester, second semester): body type, muscle tone, and posture of body regions from a front view, side view, and rear view.

Health record: brief review of nutrition, eyes and ears, and other medical findings relative to heart, lungs, skin, teeth, etc.

Measurements (first semester, second semester): anthropometric appraisals of ankle, calf, thigh, upper arm, lower arm, hips, waist, chest, height, and weight.

Physician's recommendations: ample space for medical comment regarding remedial activities.

Disposition of case: space for date and categorical check of progress.

Identifying information: name and sex of pupil, home address, phone, date of birth, school, homeroom, grade, and date entering class.

Profile Charting. Of special interest, first to parents and secondly to administrators and pupils themselves, is a visual record of progress. If improve-

TABLE XXIV
INDIVIDUAL PROFILE

GRADE 6

RACE Mixed

BOYS	Ball Bounce	Jump Rope	Jump for Height	Wall Ball	Accuracy Throw	Side Stepping	Distance Throw	Kick and Run	Climb	Chinning
Individual's Score	6.4	10	16	34	4	21	45	10.5	2	7
Number of Cases	100	100	100	100	100	100	100	100	100	100
Percentile 90	6.6	49	19.5	35	6	25	60	9.1	3	13
80	6.9	41	18.0	34	5	24	50	9.3	3	12
70	7.2	38	17.0	32	5	23	48	9.5	3	11
60	7.5	35	15.5	31	5	22	46	9.8	3	10
50	8.0	31	15.0	30	4	21	44	9.9	3	9
40	8.4	27	14.0	29	4	21	40	10.1	2	9
30	9.0	21	14.0	28	4	20	37	10.3	2	8
20	9.7	17	13.0	27	3	19	35	10.6	2	7
10	10.7	9	11.0	24	2	17	31	11.1	1	6

GIRLS	Ball Bounce	Jump Rope	Jump for Height	Wall Ball	Accuracy Throw	Side Stepping	Distance Throw	Kick and Run	Climb	Chinning
Individual's Score	10.8	40	11.5	24	3	23	23	10.7	2	8
Number of Cases	100	100	100	100	100	100	100	100	100	100
Percentile 90	7.4	57	17.0	33	4	23	34	10.0	3	11
80	8.1	51	15.5	31	3	22	31	10.2	3	10
70	8.7	48	15.0	30	2	21	26	10.4	2	9
60	9.0	46	13.5	28	1	20	25	10.5	2	8
50	9.4	44	13.0	27	1	20	23	10.8	2	7
40	9.8	43	12.0	26	0	19	22	11.0	2	7
30	10.3	40	11.5	24	0	18	21	11.2	1	6
20	11.1	35	10.0	23	0	17	19	11.5	1	5
10	12.1	28	9.0	21	0	15	17	11.8	1	3

Fig. 16-1. A Sixth-grade Profile Chart. (Courtesy Greensboro Public Schools, Greensboro, North Carolina.)

ments in personal achievement can be simply shown, considerable understanding and sympathy for the school health and physical education program will be fostered. In this connection the profile chart has value. It has already been mentioned in reference to the New York State Physical Fitness Test.

An example of how an individual profile is charted is illustrated in Figure 16-1: a profile of a sixth-grade boy is plotted against the norm performance of all sixth-grade boys in the city. Also, Figure 16-1 shows the profile of a sixth-grade girl. These profiles were sent to parents to help interpret the experimental testing program in elementary school physical education in the Greensboro, North Carolina, schools. It is not difficult to duplicate this in most schools.

Noteworthy Examples

There are, no doubt, numerous excellent programs of adapted physical education, many of which use tests to meet individual needs. Moreover, because pupils normally enjoy being tested, the tests are motivational devices as well as measuring instruments. This provides a happy marriage of instruction and evaluation.

There follow three brief examples of public school programs in three different states. The approach is varied, but the results are the same: improved human efficiency, thanks to a health and physical education experience, and a satisfied public.

Example: Philadelphia, Pennsylvania

The adapted physical education program is well organized in written form for junior and senior high school use. Considerable attention is given to body mechanics, muscle testing, and personal health needs. In one junior high school of 2,000 boys and girls there is actually a waiting list for adapted physical education because testing and measuring devices provide so much fun. The special exercise room reminds one of a penny arcade where scales turn, dials move, and lights flash on and off. Here, for example, a pupil doing an exercise for standing tall pushes against an overhead cushion, which has a dial registering pounds of pressure that can be read by the pupil as she looks ahead at the mirror. Another device is a trough arrangement for strengthening weak feet. Here the student walks with the weight to the outside of the foot, but should she walk any other way, a red light flashes on just ahead where it is easily seen. Numerous other appraisal devices of this nature make remedial activities fun as well as profitable.[8] At both junior and senior high school levels poor posture and foot deviations constitute the great majority of conditions handled. In several schools the stack of testimonial letters from satisfied parents is most impressive.

[8] See highly descriptive article by William Braverman, "Motivating Devices for Remedial Physical Education," *Journal of Health, Physical Education and Recreation,* 29:17–19, March, 1958.

Example: New Britain, Connecticut

Employing measurement to give the public a bird's-eye view of a well run physical education program, Avedisian and Bedard, of New Britain, Connecticut, worked out an organizational chart. Effective use of the chart was made at parent-teacher conferences, staff meetings, and community and civic organizations. The place of the adapted physical education program, the regular class program, and intramural and interscholastic athletics is understood as one studies the over-all organizational chart. For the first time many citizens discovered the interrelationship of physical educators, guidance personnel, school physicians, and social workers.

Here medical examinations are followed by PFI tests and pupil classification into homogeneous groups according to needs. Pupils are divided into substrength, normal, superior, and gifted categories. Adapted and regular class activities are afforded the first two groups, and leadership through intramural and interscholastic sports is stressed in the latter two groups. The director of guidance works very closely with health and physical education teachers, a practice that is most effective and could well be adopted in many more schools. It is a "work program" for the teachers, but it is well worth the effort.

Example: Moses Lake, Washington

Because of its geographical location in Washington, Moses Lake Junior High School, with an enrollment of 1,000 pupils, was selected for a pilot study in physical education for the subfit child. Originally (1956) this was an experimental program in an attempt to gain practical knowledge that would be useful to schools throughout the state.[9] Three noteworthy objectives were met in the course of the program: (1) evaluation of fitness status of several hundred junior high school pupils, (2) selection of subfit children for further study, and (3) construction of individual programs based upon pupil needs.

From the beginning this was a group approach, involving an analysis of health and guidance records, interviews with students, health council referrals, and special physical education classes. The Physical Fitness Index and the Kraus-Weber Tests were administered to all boys and girls. The Kraus-Weber battery was not used to determine how many passed or failed the exercises but to locate marked anteroposterior and lateral deviations and muscular imbalance.

After the program had been in operation about eight months, it was carefully evaluated. The results are most heartening and indicate clearly that measurement and evaluation efforts are productive in helping educators achieve their goals. Some of the significant conclusions from the Moses Lake Project are the following:

1. The group approach is important.

2. Specialists at the school were "amazed" at the many factors of common interest and how scattered information, when pooled, gave a much broader and deeper understanding of the child's problem.

3. Physical education could benefit greatly by working more closely with the medical specialists.

[9] For a detailed discussion of the experiment, see Donald K. Matthews, Virginia Shaw, and Philip Risser, "The Moses Lake Project," *Journal of Health, Physical Education and Recreation*, 29:18–19, April, 1958.

4. Although the average group was not neglected, interest centered on the low-scoring child.

5. The reporting of fitness scores to the subjects was a strong motivating device for greater performance.

6. Because of the group approach, there was continuity between the health program and total education of the child.

A Final Plea

A final note to this chapter is in the nature of a plea to conscientious health teachers and physical instructors to make the most of their unique function in education; that is, to measure all pupils scientifically and adapt the program to individual needs. The activity will be its own reward in personal satisfaction. But even more important, physical education will be redirected and will begin to achieve its proper status in western civilization. In short. survival in an already overcrowded school curriculum depends upon the ability to change and move ahead more judiciously to educational goals.

SELECTED REFERENCES

Brownell, Clifford L.: *Principles of Health Education Applied*, New York, McGraw-Hill Book Company, Inc., 1949, chap. 11.

Clarke, H. Harrison: "Schools of Physical Education and the Training of Physical Medicine and Rehabilitation Specialists," *Journal of the Association of Physical and Mental Rehabilitation*, 7:2 6, March-April, 1953.

Daniels, Arthur S.: *Adapted Physical Education*, New York, Harper & Brothers, 1954.

Dunkelberg, James G., and Gene A Logan: "Let the Doctor Recommend Adapted Physical Education," *Journal of Health, Physical Education and Recreation*, 29:28–29, May-June, 1958.

Mueller, Grover, and Dorothy P. McQueen: "Individual Remedial Physical Education," *The Physical Educator*, 9:88 91, October, 1952.

"School Health Services," *Report of Joint Committee on Health Problems in Education of the National Education Association and the American Medical Association*, Washington, National Education Association-American Medical Association, 1953, chap. 8.

CHAPTER 17

Administration of Evaluation Programs

The ultimate success of any educational endeavor reflects the stimulation, support, and zeal of the administrator. Although the alert teacher of health or physical education is a real asset in the health program, the value of a school administrator who is sympathetic to the goals of school health is highly significant.

In small systems the school administrator works directly with his health and physical education teachers. In larger systems much of his responsibility is delegated to the director of health and physical education. In either case, the administrator "sets the stage" for the cooperative work (team approach) of physicians, dentists, nurses, psychologists, dental hygienists, physical educators, and health teachers. He appraises the efforts of the school custodial staff as they maintain safe and sanitary buildings. The health-oriented administrator promotes in-service education for medical personnel, physical educators, and classroom teachers alike. He maintains lines of communication between school and community agencies for several reasons, including securing adequate treatment for children in special need or obtaining funds for program evaluation. He arranges for periodic nurse-teacher conferences and meetings of health councils. He appraises school safety practices periodically. He supports supervisors and teachers in their quest for improved methods of instruction in health and physical education. He also provides the kind of leadership necessary to carry on research relative to such health items as school absences, pupil health problems, and various measures of individual pupil and program efficiency. In short, the administrator has as much to do with evaluation as with goals, curriculum, and instructional methods and materials.

Administration in Action

Any worthwhile undertaking has an underlying philosophy, which substantiates whatever is planned for in the scheme of things to come. Educators have philosophized about the world they hope to make. Each succeeding generation has talked about the great deeds left to be done, the new frontiers to conquer, and the utopias yet to be fashioned. Each age has had its great administrators who were remembered as great because they did more than dream; they measured their progress carefully and got things done.

In administering a measurement program in health and physical educa-

tion one must plan ahead so effort will not be haphazard or incomplete. To ensure that administrative appraisals are wisely conceived, carefully planned, efficiently conducted, and properly followed up, nine specific steps are recommended (see Figure 17-1).

Many excellent examples of school systems that actually have their total school program built around the results of health and physical education evaluation are in operation across the country. Communities such as Greensboro, North Carolina; Needham, Massachusetts; New Britain, Connecticut; Philadelphia, Pennsylvania; and Los Angeles, California, have had strong administrative support as well as competent leadership at the teacher-pupil level for a number of years. In fact, sound appraisal programs do not often

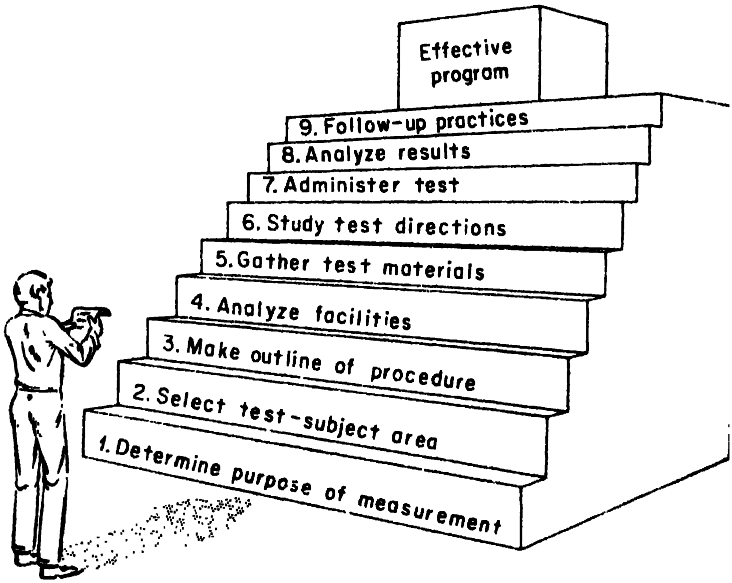


Fig. 17-1. Steps in Administering the Measurement Program.

come about in a year or two. They arise out of a sincere effort to apply the nine steps:

1. Determine the Purpose of Measurement. Simply stated, what are you seeking and can you find out by taking time from the instructional program?

2. Select Test for Subject Area. Isolate the health or physical education unit to be measured. Delimit the problem. No tests measure the whole program; select the best test or tests in any one area. Subject it to scrutiny by relating it to the criteria of proper measures.

3. Make an Outline of the Procedure. Is this to be an individual or a group test? If it is a group test, how many persons are to be tested in a given period of time? The size of the group to be tested depends upon the nature of the test. The organization of pupils depends, more specifically, on the manner in which the available space is utilized, the arrangement of the testing stations, and the provision of satisfactory leadership.

4. Analyze the Facilities. The previous step cannot be completed until this step has been accomplished. In fact, the finest test available may prove useless if school facilities are inadequate. Space planning to eliminate congestion and confusion ahead of time is efficient administrative practice. Most of the carefully developed tests state specifically what is required in the way of space. For example, the manual for the New York State Physical Fitness Test reads as follows:

> One posture testing station requires an area of approximately 5 × 15 feet. One or more of these stations, depending upon the availability of qualified raters, may be set up in the gymnasium or in any other convenient place such as the locker room or the school nurse's office.

Also, in the same test manual, other items, such as the speed, accuracy, and agility measures, are carefully sketched showing the arrangement of testing stations.

5. Gather the Test Materials. Everything needed from special instruments and sporting equipment to pencils and paper should be on hand at the time of the test.

6. Study the Test Directions. The nature of the test questions and requirements should be clear to the person doing the testing. Moreover, the persons to be tested should also be told the nature and purpose of the testing activity. All too often tests are hastily administered to pupils who do not understand what is to be accomplished; the result is a low level of personal motivation. Properly oriented pupils know the purpose of the test, have either tried part of it or seen it demonstrated, know the layout of the testing stations, know their responsibility for scoring and recording data, and are familiar with the dates, times, places, and uniforms for the testing.

7. Administer the Test. Since test results are usually interpreted on the basis of test norms, it is essential that the specific directions and procedures be followed exactly. Obviously, if some pupils are given a longer time to perform a test item, or if they are given special help or consideration during the test, their scores will not be comparable with those of pupils who did not have such special advantages.

8. Analyze the Results. When the data are collected, what does it mean? How do the pupils in this school compare with others elsewhere? What is the present status? Has there been some improvement? For example, in one Stratford, Connecticut, junior high school a large segment of the male population (51.9 per cent) failed to pass the Kraus-Weber Tests. After an adapted program of systematic muscle training for five months in physical education classes, the failure rate was reduced to 21.7 per cent, a considerable improvement. The results are more enlightening when analyzed further. In this study it was found that muscular fitness begins in the home and that pupils need a chance to do more walking and get more physical activity in general.

9. Follow-up Practices. This is the top step, the place to do something about the findings. Unfortunately, far too many measurement programs slow down at this point. Follow-up practices designed to bring about pupil improvements frequently require careful administrative thought and a team aproach to make the desired changes. In Greensboro, North Carolina, for example, pupils are individually motivated to improve their performance by studying the norm tables with the median performance of other children of their own age. Also individual profile charts are constructed for the parents to motivate them to help bring about improvements.

There are three follow-up questions to be answered in every school. These can be checked by a person designated by the administrator, probably the nurse.[1]

1. Are children who need professional attention being found?
2. Are children who are referred for professional attention receiving it?
3. What do pupils and parents learn about health as a result of health appraisal and counseling?

That administrative measures keep the school on a sound footing is well demonstrated by a study of the Los Angeles city schools. Here a number of administrative surveys have been carried out of school sanitation and related areas. For example, the policy to give audiometer tests was appraised. In order to determine the adequacy of this policy, a study was made in four elementary schools.[2] It was found that at least 50 per cent of the pupils with significant hearing loss (more than 15 decibels loss in either ear) were not being identified in any one school year by the policy of testing pupils in two selected grades. The facts elicited by this study were helpful in showing the need for additional school audiometrists and otologists. These were later added to the school staff.

Administrative Weaknesses

A number of administrative practices are weak in application or missing altogether in some schools The team approach, so often referred to, is nonexistent, and physical educators hardly know the name of the school physician.

In analyzing school health service practices, Neilson and Irwin showed that different practices are related to different administration plans.[3] Moreover, there was a lack of coordination between school health services,

[1] "School Health Services," *Report of Joint Committee on Health Problems in Education of the National Education Association and the American Medical Association,* Washington, National Education Association–American Medical Association, 1953, p. 405.

[2] C. Morley Sellery and Blanche G. Bobbit, "Evaluation of Health Education and Health Service in the Los Angeles City Schools," *Journal of School Health,* 30:113–118, March, 1960.

[3] Elizabeth A. Neilson and Leslie W. Irwin, "Analytical Study of School Health Service Practices in the United States," *The Research Quarterly,* 29:417–458, December, 1958.

physical education, and health instruction. There was also a noticeable lack of parent conferences, insufficient community aid for children who could not afford to pay for remedial services, insufficient number of planned conferences between individual teachers and nurses for every child, and a lack of adequately trained personnel.

Administrative Measures

There are several useful appraisal instruments designed to provide the administrator with a somewhat broad view of the health and physical education program. Some are standardized; many are used locally or employed experimentally. These are of value, however, when used as guides in the over-all appraisal of the program. Some of the health program appraisal devices have already been referred to in Chapter 4.

The publication of the Texas State Board of Education, *A Checklist Appraising the School Health Program* (1955), is one of the most current and thorough measures available for public school use.[4] The total health education effort is reviewed in the light of specific standards.

The *Indiana Physical Education Score Card for Elementary and Secondary Schools* (1958) was developed to improve the quality of physical education programs.[5] It is primarily a self-evaluation instrument, which encourages follow-up studies. The score is divided into four areas, which are weighted in the following manner:

Area	*Per cent of total value*
1. Class management and instruction	40
2. Program activities	30
3. Facilities and equipment	20
4. Administration	10

Because of differences in programs for boys and girls and for elementary and secondary pupils, there is a separate score card for secondary school boys, secondary school girls, lower elementary school grades one, two, and three, and upper elementary school grades four, five, and six. Each group is scored in each area after several questions have been answered for each area.

Example

Administration

Secondary Boys

1. Pupil-teacher ratio for physical education

35 or less pupils	6
36–45 pupils	4
46–55 pupils	2
56 and above pupils	0

[4] Copies can be obtained from the Texas Education Agency, Austin, Tex.

[5] Copies can be obtained from the Division of Health and Physical Education, 1330 W. Michigan St., Indianapolis, Ind.

2. Days per week of physical education

4–5 days	6
2–3 days	4
1 day	2
0 days	0

6. Substitutions for required physical education

No substitution permitted	9
Athletics during season	6
Band and/or ROTC	3
Any other substitution	0

7. Grading system for physical education

Same as other subjects	9
Pass or fail—where other subjects use letter grades	6
No grade	0

Much credit should be given to the Indiana physical education score cards. They are easy to use and can effectively point to weaknesses within the school system. Moreover, this instrument, when employed as a self-evaluation device, is likely to awaken the administrator to acceptable policies more quickly than almost any other procedure. Herein lies its strength.

The *LaPorte Health and Physical Education Score Card* (I, II) has probably been the most widely accepted evaluation instrument for administrative appraisals.[6] It is intended as a measuring device not only for the physical education program but for the general health, safety, and recreation provisions of an entire school. The rating is to be made by the school principal himself or by his official representative assisted by the physical education instructor. The card is set up in such a way that it centers attention upon the characteristics of a good program. It was validated after a twenty-three-year study by the Committee on Curriculum Research, of the College Physical Education Association. There are ten areas of concentration, consisting in ten elements each.

Area of concentration	*Possible score*
I. Program of activities	30
II. Outdoor areas	30
III. Indoor areas	30
IV. Locker and shower areas	30
V. Swimming pool	30
VI. Supplies and equipment	30
VII. Medical examinations and health service	30
VIII. Modified individual (corrective) activities	30
IX. Organization and administration of class programs	30
X. Administration of intramural and interscholastic athletics	30
Total possible score	300

[6] William R. LaPorte, *Health and Physical Education Score Card II* (*Secondary Schools*), Los Angeles, Calif.; Parker and Company, 1951. Score Card I is designed for elementary school use.

Each element can be rated from 1 to 30 points; a perfect score is 300 points. If 200 points are scored, the program is considered fair to good; if only 100 points are scored, the program is a poor one. If desired, the scores can be reduced to percentages. Moreover, a profile chart can be accomplished on a form provided. Reliability coefficient for the total score card, based on judges' ratings, was .98 in a study carried on by Dana Clark.[7] Using the LaPorte score card Clark discovered that the quality of health and physical education was related to size of the school. The larger the schools, the better the program. The best provisions were in schools enrolling at least 200 pupils in the upper three high school grades. Similar findings occurred in Ohio, where the score card was used to show that accreditation and school size have considerably more influence upon the quality of the health and physical education program than geographic area, town size, or consolidation.[8]

One of the largest surveys employing score card II was carried out on 135 North Carolina high schools as part of a physical education study.[9] Scores ranged from 32 to 201, with a mean of 91.38 and a median of 76. As in the New York and Ohio studies, school size was significantly related to favorable scores. Studies such as these serve to illustrate the practical nature of the LaPorte Health and Physical Education Score Card.

A *Rating Card for Athletic Directors* was fashioned by Kelliher and used by Rice to evaluate thirty-six items related to the athletic program.[10] The athletic director rates (1) financial soundness, (2) organization of the department, (3) professional status of the staff, (4) well-being of the staff, (5) well-being of the students, (6) public relations, and (7) care of property and equipment.

In order to measure the quality of men's physical education in higher education in New York State, Phillips developed a questionnaire.[11] It was based on a study of the needs of college students plus broad principles of physical education. The questionnaire covers general information, administration and supervision, institutional requirements, course requirements, attendance, staff, program, classification of students, budget, facilities, and equipment.

A *Check List for Physical Education,* distributed by the New York State Education Department for use by school officials, consists in a number of

[7] Dana E. Clark, "Quality of Administrative Provisions for Physical Education in New York State," *The Research Quarterly,* 25:269–276, October, 1954.

[8] Lawrence E. Owens, "Health and Physical Education Program for Boys in Ohio High Schools," *The Research Quarterly,* 26:461–469, December, 1955.

[9] Nathan Taylor Dodson, "Evaluation of Physical Education in North Carolina High Schools," *The Research Quarterly,* 27:243–244, May, 1956.

[10] Harry M. Rice, "Teamwork," *Journal of Health, Physical Education and Recreation,"* 30:30–31, December, 1959.

[11] Byron M. Phillips, "Evaluation of Men's Physical Education Service Programs in Higher Education in New York State," *The Research Quarterly,* 26:185–196, May, 1955; questionnaire can be obtained from the author at 14 Pleasant St., Canton, N.Y.

"yes" or "no" statements involving the items of program activities, program planning, evaluating, scheduling, personnel, facilities, safety and sanitation, and budget.[12] It has had wide use in New York and adjacent states and can be obtained free of charge.

Several years ago Van Dalen prepared a useful check sheet, *Functional Evaluation of the Health and Physical Education Curriculum.*[13] A number of the recognized goals of education and health and physical education are listed under seven different headings. The degree to which an opportunity exists for striving toward these goals is rated as *N* (no opportunity), *S* (some opportunity), *C* (considerable opportunity). Space is provided to note just where the experience takes place: in the health program, gymnasium program, intramural program, remedial program, or interscholastic program. The questions are scholarly and require some contemplation before the person doing the rating makes a mark on the score sheet.

Acting on the premise that school health administrators oftentimes are guilty of endorsing programs based on many instances of apparent accomplishments or on procedures handed down from the past, the Los Angeles city schools health service personnel decided to evaluate their programs. The result was the creation of three short and highly practical forms: (1) *Appraising the Elementary School Health Program,* (2) *Appraising the Secondary School Health Program,* and (3) *Appraising the Junior College Health Program.*[14] The questions asked have depth and are very revealing. Here is one from the secondary school form:

17. Approximately what per cent of the following pupils had a health examination during the school year 19____ to 19____?

	Approximately		
	50%	75%	95–100%
Pupils in Grade 7	____	____	____
Pupils in Grade 10	____	____	____
All new pupils to the school	____	____	____
Pupils referred by teachers	____	____	____
Rechecks of special cases	____	____	____

Appraisal forms such as these cannot help but reveal program weaknesses, which by and large are susceptible of improvement.

Working at the University of Illinois, Jackson developed a battery of questions designed to scrutinize the health education program.[15] These were

[12] *Check List for Physical Education,* Albany, N.Y., The New York State Education Department, Division of Health and Physical Education, 1952.

[13] D. B. Van Dalen, "Functional Evaluation of the Health and Physical Education Curriculum," *The Physical Educator,* 6:1–3, March, 1959.

[14] Copies can be obtained from the Los Angeles City School District, Auxiliary Services Division, Health Services Branch, Los Angeles, Calif.

[15] C. O. Jackson, "Let's Rate Your Health Education Program," *Journal of Health, Physical Education and Recreation,* 26:29–30, September, 1955.

tried out and prepared in a check list so a school health council or committee might make use of the results. The four areas covered are administrative, curricular, instructional, and community. This check list, which appears in the cited reference, is scored "yes" (5 points), "to some extent" (3 points), or "no" (0 points). Each question is made especially clear by an elaboratory statement following it. Here is a sample question from the instructional area:

Yes	No	To Some Extent	
			1. *Is health education considered important enough to be a required subject, with credit, for all freshmen, and either juniors or seniors?* Since health is a foremost objective of education, organized emphasis in this area should not be left to chance, or made elective. Teaching health education as a "rainy day" subject in physical education is indefensible. Rather, such a course should be planned to meet the needs of the pupils, and required, with credit for the equivalent of two semesters.

One of the most extensive efforts to analyze and compare school health practices was made by Neilson and Irwin.[16] A special inquiry form, the bulk of which appears in the cited reference, was designed to evaluate school health services and administrative practices. It was validated by a jury of experts: five from health education, five from public health, and five from the field of medicine. The final form consists in sixty-one categories, with 620 options arranged under six major divisions. Neilson and Irwin used the form to check on health service practices in communities of varying size, but it can be used to measure practices in any community.

Additional administrative measures applicable to health education practices are referred to briefly in Chapter 4.

An excellent check list that the administrator can use to survey a given school athletic situation was prepared by the Educational Policies Commission.[17] This *Checklist on School Athletics* was developed primarily to help school officials appraise and improve their athletic policies and programs. A copy appears in Appendix B. Each of the 100 questions is answered simply with a yes (Y), no (N), or unknown or uncertain (U). Each question is also answered with a check mark under agree (A), disagree (D), or uncertain (U). The degree to which a school administrator answers questions with a yes (Y) indicates that practices are in keeping with the policies recommended by the Educational Policies Commission. The degree to which a school administrator answers questions in the agree (A) column measures the extent to which his opinion coincides with the opinion of the Commission.

[16] Elizabeth A. Neilson and Leslie W. Irwin, *op. cit.*

[17] *School Athletics: Problems and Policies,* The Educational Policies Commission, Washington, 1953.

Reporting Appraisal Findings

The usefulness of test results is directly related to the administrative handling of test records and scores. Scoring blanks should be carefully thought out so they contain useful information. The scores obtained should be meaningful, not only to the instructor, but to the student and parents. Bringing the test results to the attention of the children is practical in terms of instruction. Bringing the test results to the attention of the school administrators and the public tends to ensure a sound school-community relationship.

Tables, graphs, profile charts, and form letters all reflect evaluation information. In New York State the suggested practice is that the parents be sent a rather detailed form letter along with the pupil's cumulative record form for the New York State Physical Fitness Test. This clears up simple questions and encourages individual parent-teacher conferences.

Graphs work well with both student and adults. It should be well proportioned, suit the data pictured, tell the truth, and make clear beyond doubt what it represents. The advantage, therefore, in graphic presentations is that data can be shown in such a way that their significance can be understood without intensive study and frequently at a single glance. There are bar graphs, histograms, line graphs, percentile graphs, and profile graphs. All have a purpose.

If at all possible, pupils should know the results of any test they take. Keeping them uninformed serves no purpose, except to reduce individual motivation on succeeding examinations. Class announcements, bulletin board displays, etc., are helpful administrative techniques in an evaluation program.

Research in Health and Physical Education

Research is an administrative responsibility in every school. Good teaching is supplemented by a certain degree of experimentation and program modification. Not every school can arrange to carry on significant research studies, but a positive research viewpoint should permeate every health education and physical education effort. One should ask, "How may I improve my curriculum?" "What curricular information may I obtain from the professional literature?" "How much curriculum experimentation may I arrange in my own school?"

The list of unsolved problems in health education, coupled with a similar list for physical education and recreation, would easily fill a small textbook. There is no need, therefore, to list them all here. It should be said, however, that research on these problems is very much a part of evaluation. Good program appraisal, accomplished in the light of proper measurement techniques and leading to change in the curriculum and the methods of teaching, is a form of practical research especially needed in today's schools. In

modern society there are many changes, and the school is expected to change too, not haphazardly but on a scientifically sound basis. Thus, administrators work to improve their policies, offerings, and general procedures through continuous evaluation and research. There is no other way to ensure orderly progress toward the aims of education, nor is there any other way so effective in advancing both the profession and the professional worker.

The Redirection of Health and Physical Education

Never before in modern history has the case for health education and physical education been so well supported by research in medicine, sociology and anthropology, psychology, and educational philosophy. The way is open and the path is clear ahead for the kind of an educational program that will fulfill the human need for an abundance of physical, mental, social, and spiritual health.

To health and physical educators the opportunity exists now more than ever before to capitalize on medical findings and demonstrate in community after community that the youth of the land are better able to serve themselves and mankind as a result of healthful activities in school. This, of course, must be proved. In short, it is time for a redirection in health and physical education. It is time for much more scientific measurement of physical and social efficiency. It is time for programs adapted to individual needs and for adapted physical education classes. It is time for the "team approach," where the nurse, physician, and physical education teacher work cooperatively with all extremes of physical and social well-being. It is time that physical educators concentrated on the most valid tests available, even if they take a little more effort and require more knowledge to interpret. The use of valid and precise measuring instruments is an absolute prerequisite for gaining recognition as a profession, whereas haphazard testing spells mediocrity, ineffectuality, and worse. The cry for simplicity in testing may be a serious mistake. Proper measures, supported by scientific research, gain the respect of other professions and the public alike. Simply because some physical educators are not qualified to evaluate properly, there is no excuse for not recommending better appraisal devices.

The profession that is effective first pauses for redirection, sets its sights high, and then moves vigorously ahead. It is time for such redirection in health and physical education. It is time now for the profession to mature and to rise in respect and dignity.

SELECTED REFERENCES

American Association for Health, Physical Education, and Recreation, *Research Methods in Health, Physical Education and Recreation,* 2d ed. Washington, National Education Association, 1959.

Brownell, Clifford L.: *Principles of Health Education Applied,* New York, McGraw-Hill Book Company, Inc., 1949, chaps. 6 and 11.

Esslinger, Arthur A.: "Yesterday, Today and Tomorrow," *Journal of Health, Physical Education and Recreation,* 30:19–20, September, 1959.

Landis, Paul E., and P. C. Bechel: *Ohio High School Standards for Health Physical Education, Recreation and Safety,* Columbus, Ohio, Ohio State Department of Education, 1949.

National Education Association and American Medical Association, *Suggested School Health Policies,* National Education Association, Washington, 1956.

Wise, John E.: "Evaluating Research," *School and Society,* 85:354–355, Nov. 23, 1957.

CHAPTER 18

Interpretation of Data

Prove all things; hold fast that which is good.
Thessalonians I:5:21

It is not the purpose of this chapter to do justice to the topic of statistics in education. The chief concern is that the student of measurement in health education and physical education understands how test scores and similar data can be interpreted. All too often erroneous conclusions are drawn from test results, or highly positive statements are made when data indicate only slightly positive findings. This, of course, is the old practice of using half-truths to generalize. It is illustrated daily by expressions on radio and television and in the classroom, where it is stated that "most people" do this or do that or eat this or eat that. Frequently such utterances are backed by subjective opinion or at best by a study of averages where 51 per cent acted in a certain way. If slightly over half the cases react favorably to an item, it is not right to imply that those who do not act likewise are in some manner atypical or even abnormal. For this reason, statistics are viewed with some skepticism; Disraeli said, "There are lies, damn lies, and statistics."

Elementary statistics is concerned with common statistical terms and formulas without regard to mathematical derivation. Specifically *statistics have to do with numerical facts of measurement, collected systematically and classified so one is able to abstract from the data the utmost in the way of significant, dependable conclusions.*

The need for statistics was nicely expressed by Lord Kelvin, in 1883, when he said:

> I often say that when you can measure what you are speaking about and express it in numbers you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of meagre and unsatisfactory kind: it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science.

Care in Using Statistics

One must be careful in using statistics to plan statistical studies so results can be easily and readily interpreted. There are too many persons in health

and physical education who will not subscribe to a professional research magazine because they claim they cannot understand it. Also statistical material should be such that it can be easily compared with the work of others in the same area.

The need for taking a close look at statistics and properly interpreting them is well illustrated by the fact that one may be scared to death by train statistics. In one year, for example, 4,712 deaths were attributed to railroad transportation. The implication might be to stay off trains. On further analysis of the figures, it is discovered that over half the deaths resulted from automobiles colliding with trains at crossings, and most of the rest were persons "riding the rods." Only 132 were passenger deaths.

Another point involves the relationship between observations and conclusions. Among some native tribes, observation over the centuries taught them that persons in good health usually had lice. Very often sick persons did not have them. This was an accurate and sound observation, for lice often leave a body with fever. The logical conclusion, however, that lice make a person healthy and that everyone should have them, is erroneous. Here cause and effect have been altogether confused. Had some elementary statistics been applied to the number of cases having lice, the implications derived from observation might have been quite different.

In short, statistics give meaning to the results of observation and the more complex processes of measurement.

Grouping of Data

When the testing is completed and a large number of test scores are on hand, it is necessary to consider just what these scores mean. Unless there is a norm table or some other standard point of reference to refer to, an individual test score may mean very little. Information on how one subject compares with his group or how his group compares with others can be obtained only when the raw scores are assembled in some logical fashion. Such a logical arrangement is called a *frequency distribution.*

Suppose a class made the following scores on a health knowledge test: 33, 30, 31, 37, 30, 33, 34, 35, 33, 36, 33, 34, 37, 29, 34, 36, 39, 31, 34, 30, 33, 37, 32, 38, 36, 27, 33, 31, 34, 32, 34, 31, 32, 36, 29, 33, 35, 31, 35, 32. By distributing these scores into a frequency table, such as Table 18-1, they are readily understood.

From the table it is evident that one person scored 27; no one made a score of 28; three persons scored 30, etc. The scores actually go halfway to the next number; that is, the score 27 goes from 26.50 to 27.49; 28 goes from 27.50 to 28.49.

Usually the scores are grouped into intervals or classes. In the frequency table these are shown as step intervals (i). Each interval has its upper and lower limits. In the above illustration there are thirteen intervals, the size of which is 1. If score limits happened to be 26.50 to 29.49, 29.50 to 32.49,

Table 18-1
Frequency Table of Scores on a Health Knowledge Test

Score	Tallies (N = 40)*	f†
39	1	1
38	1	1
37	111	3
36	1111	4
35	111	3
34	11111 1	6
33	11111 11	7
32	1111	4
31	11111	5
30	111	3
29	11	2
28		0
27	1	1

*N = number of scores or sum of fs
† f = frequency with which each score occurs

32.50 to 35.49, etc., the size of the step interval would be 3. In selecting the size of the interval, subtract the poorest score from the best score. This gives the range. If there was a range of 47 score points, a 3-step interval would prove satisfactory and would yield approximately sixteen intervals (47/3). Table 18-2 has a range of 78 score points and a step interval of 5. As a

Table 18-2
Frequency Distribution on a Sports Knowledge Test

Score	Mid-points	Tallies (N = 50)	f
89.5–94.4	92	1	1
84.5–89.4	87	11	2
79.5–84.4	82	111	3
74.5–79.4	77	111	3
69.5–74.4	72	11	2
64.5–69.4	67	11	2
59.5–64.4	62	111	3
54.5–59.4	57	1111	4
49.5–54.4	52	1111	4
44.5–49.4	47	11111 11	7
39.5–44.4	42	1111	4
34.5–39.4	42	1111	4
29.5–34.4	32	111	3
24.5–29.4	27	111	3
19.5–24.4	22	11	2
14.5–19.4	17	11	2
9.5–14.4	12	1	1

general rule, the number of step intervals is not less than 10 or more than 20; the preferred size of the interval is 2, 3, 5, 10, and 20.[1]

Measures of Central Tendency

When the scores have been arranged into a satisfactory distribution it is desirable to find one score that represents all the scores made by the group. Such a score indicates the *central tendency* of the group and is expressed by the mode, median, and mean. The *mode* is the score appearing most frequently. In Table 18-2 above, the most typical score falls in the interval with the frequency of 7. It is assumed that the mode lies somewhere between 44.5 and 49.4; that is, 47. The mode, at best, is a crude measure of central tendency, used where a rough estimate is satisfactory.

The *median* is employed where the exact mid-point of a distribution is desired. It is the score above which and below which lie an equal number of scores. A crude median, therefore, is obtained by simply dividing the total number of scores by 2. If there were sixty-two cases, the median would be at the thirty-first case. In a frequency distribution of scores the median is found by interpolation. If $N/2$ is applied to Table 18-2, there are exactly twenty-five scores on either side of the median. By adding up the f column, we have nineteen cases when the score 44.4 is reached. Thus we need six out of the seven cases, or $6/7$, of the next class interval, which consists of 5 points.

$$\frac{6}{7} \times 5 = \frac{30}{7} = 4.2$$

This means we have to go 4.2 points beyond 44.4, so we add 4.2 to 44.4, which equals 48.6. This is the median score. The median is often used when a distribution of scores is out of balance, by being concentrated at either the top or bottom end of the distribution. Such a distribution is considered to be *skewed*. In a perfectly symmetrical distribution the mode would be as good a measure of central tendency as the median. Usually, when there are only a few scores in a frequency distribution, they are out of balance, and it is here that the median provides a more accurate measure of central tendency.

The median is, of course, the fiftieth percentile. Being at the mid-point of the distribution, 50 per cent of the scores are above the fiftieth percentile and 50 per cent are below.

The *mean* is the average and best measure of central tendency. It is used when greater reliability is wanted, when other computations are to follow, and when the frequency distribution is symmetrical about the center. When scores have not been put into a frequency distribution (ungrouped data), the mean is easily determined by adding all the scores and dividing by the number of cases (N). This is sometimes called the *arithmetic mean.*

[1] J. P. Guilford, *Fundamental Statistics in Psychology and Education,* 3d ed., New York, McGraw-Hill Book Company, Inc., 1956, chap. 2.

Where large numbers of scores are involved, however, they are grouped in a frequency table and the short method of finding the mean is employed. In calculating the mean, a guess or estimate is made as to the mid-point of the distribution. This guess is then corrected in terms of deviations from the mid-point. The following formula is used:

$$M = M' + i\frac{(\Sigma fd)}{N}$$

where M = mean
M' = guessed mean[2]
i = step interval size
d = deviation[3]
Σfd = sum of $f \times d$

Table 18-3 is similar to Table 18-2 in that it has the same basic scores and frequencies and is an effort to establish the mean.

Table 18-3
Calculation of the Mean*

Scores	Mid-point	f	d	fd	
89.5–94.4	92	1	8	8	
84.5–89.4	87	2	7	14	
79.5–84.4	82	3	6	18	
74.5–79.4	77	3	5	15	
69.5–74.4	72	2	4	8	
64.5–69.4	67	2	3	6	
59.5–64.4	62	3	2	6	
54.5–59.4	57	4	1	4	+79
49.5–54.4	52	4	0	0	
44.5–49.4	47	7	−1	−7	
39.5–44.4	42	4	−2	−8	
34.5–39.4	42	4	−3	−12	
29.5–34.4	32	3	−4	−12	
24.5–29.4	27	3	−5	−15	
19.5–24.4	22	2	−6	−12	
14.5–19.4	17	2	−7	−14	
9.5–14.4	12	1	−8	−8	
					−88

* $N = 50$.

$$\begin{aligned} M &= M' + i\left(\frac{\Sigma fd}{N}\right) \\ &= 52 + 5\left(\frac{-9}{50}\right) \\ &= 52 - .9 \\ &= 51.1 \end{aligned}$$

[2] Sometimes written AM (assumed mean).
[3] Sometimes x^1 is used to express deviations.

In arriving at the mean in Table 18-3, the following steps should be taken:

1. Select the mid-point of some interval near the middle of the distribution as the guessed mean (in the illustration M', the guessed mean is 52).

2. The "d" column represents deviations both ways from the guessed mean. Block out the guessed mean interval by drawing horizontal lines above and below, as in Table 18-3. Number the intervals above the guessed mean as positive deviations and those below the guessed mean as minus or negative deviation.

3. Multiply the frequency times its deviations (fd).

4. Total separately the plus fds and the minus fds, and find the algebraic sum (the difference) of the two. In Table 18-3 the plus fds add up to $+79$. The minus fds add up to -88. The difference, Σfd, is -9.

5. Divide Σfd by the total number of cases (N) and multiply this by the size of the step interval (5). This is the correction procedure that will add or subtract from the guessed mean and give a true mean figure. In the above example $i(\Sigma fd/N) = 5(-9/50) = -.9$, which is subtracted from the guessed mean ($52 - .9$) to yield an actual mean of 51.1. If Σfd had been positive, $i\ (\Sigma fd/N)$ would have been added to the guessed mean.

The mean is used the most in statistical procedures. Here each score has equal weight for its full value; that is, the central tendency of a group of scores is affected by the pull of individual scores at each end of the distribution. Where the median indicates only the center of the distribution, the mean reflects the pull of extreme scores.

Percentiles, Quartiles, and Deciles

Both the student and the instructor are interested in the relative position of an individual score, especially when a large number of students are involved in an important examination.

Percentile scores are simply scores that break down the total distribution into segments. They give meaning to scores that otherwise might be largely meaningless. For example, if a student learns that he is at the tenth percentile in a particular test, he should understand that 90 per cent of his colleagues did better than he on the test. If he were informed that he had a score of 42, he might have little idea as to whether he did very well, fair, or poor until he checked the percentile level. Percentiles are computed the same as the median. Thus the thirtieth percentile score in Table 18-2 would fall at the fifteenth score from the bottom (39.5), because 30 per cent of fifty scores is 15. Likewise a student scoring above the ninetieth percentile level would have a test score above 72.

When the distribution is divided into quarters, the four segments are referred to as *quartiles*. On Table 18-2 a student scoring above 65 would be in the upper 25 per cent or upper quartile (Q_4). A student scoring below 37 would be in the lower 25 per cent or lower quartile (Q_1).

Most percentile tables go by units of ten, called *deciles*. It is not difficult to set up this kind of table from raw data. To establish a percentile scale for the scores in Table 18-2, one begins at the bottom. As there are fifty pupils, the first decile, or tenth percentile, calls for 10 per cent of fifty pupils, or five pupils; that is, the fifth pupil is at the tenth percentile. By looking at Table 18-2, it is discovered that the fifth pupil made a mid-point score of 22, so 22 becomes the level for the tenth percentile. In like manner, the twentieth percentile falls at the score of 32. In this manner the table is completed.

Sometimes the test scores from previous examinations are compared with present scores, or several groups of scores may be compared. For example, the results of initial physical fitness testing (PFI) in one junior high school indicated that the eighth-grade boys scored higher than either the seventh or ninth graders[4]:

Distribution	Seventh grade	Eighth grade	Ninth grade
High score	165	171	160
P_{90}	124	130	128
Q_3	111	118	116
Median	98	100	100
Q_1	89	88	82
P_{10}	75	79	69
Low score	51	53	51

Here, percentiles and quartile levels are used together, and all three sets of scores can be compared on an equal basis.

Measures of Variability

Averages are notorious for what they conceal. One speaks of the average salary of school teachers and then overlooks the fact that most such teachers do not get this average salary. Here the "pull" of several high salaries offsets a greater number of much lower ones causing the average to be frequently misleading.

Averages, medians, and modes attempt to express the whole range of measurement results by one figure. This is the central tendency of the distribution. Although it gives the tester an idea of where most students score, it does not show the scatter, or spread, of the scores, that is, the variability within the distribution. Central tendency is concerned with the group as a whole, but variability is concerned with the differences in individual scores.

The *range of scores* is a rough measure of variability; it shows only the distance between the highest and lowest scores. This could be deceiving. For example, if a range of scores extended from 3 to 92, we might infer that the

[4] Actual scores are from Robert McCollum, Wilson Junior High School, Eugene, Ore.

scores are evenly or symmetrically distributed between the two points. Actually this might not be the case. The scores could be concentrated near the top of the range with only three scores below 50, or vice versa. In order to give a more adequate indication of the spread of the scores, several measures of deviation or variability are used.

The *quartile deviation* (Q) or semi-interquartile range is a measure of variability that expresses half the range of the middle half of the scores in a frequency distribution. In other words, it is half the range between the twenty-fifth and seventy-fifth percentiles (Q_1 and Q_3).

$$Q = \frac{Q_3 - Q_1}{2}$$

This is used only with the median, and it shows the dispersion of sccres away from the median. It has its limitations because it does not consider the total range of scores. It is comparable to the *probable error,* a measure that

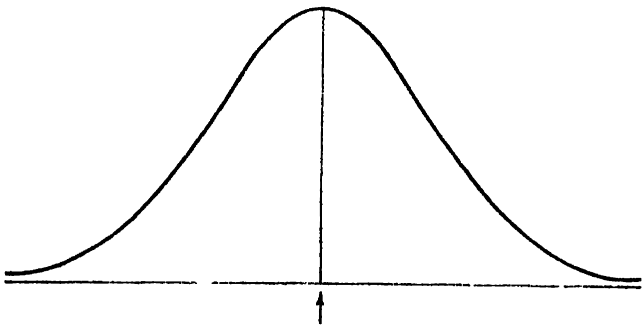


Fig. 18-1. The Normal Probability Curve.

also marks the limits of the middle 50 per cent of the scores but can only be used with a perfectly symmetrical distribution.

The *standard deviation* is by far the most reliable and most frequently used measure of variability. It is designated by the Greek sigma (σ). In the normal distribution of scores the standard deviation is defined roughly as approximately the middle two-thirds of the distribution (68.26 per cent). More specifically, it is the square root of the mean of the squared deviations taken from the mean of the distribution.

The perfectly symmetrical distribution of scores on a frequency table would form a normal distribution of scores; that is, the scores would be distributed equally from one end of the range to the other, with a concentration toward the center. This is the normal or bell-shaped curve (see Figure 18-1). It is often called the *normal probability curve,* because when a large number of cases, scores, or traits are plotted on a graph, they tend to approach the bell-shaped form. The greater the number of cases, the more this is apt

to be true. In short, where the laws of chance operate, there would be as many scores on one side of the distribution as there would be on the other.

A normal curve may be built experimentally from some experimental data where a trait or occurrence appears as often as pure chance permits. Penny tossing is the classic illustration. If, for example, one tossed three pennies eight times, the chances are that the number of "heads" and "tails" would be in proportion indicated by raising $(H + T)$ to the third power.

$$(H + T)^3 = H^3 + 3H^2T + 3HT^2 + T^3$$

Thus, in eight throws one should expect one throw of three heads, three throws of two heads and one tail, three throws of one head and two tails, and one throw of three tails. It is not likely that chance would operate in such a way as to make the above estimate faulty. After all, very little affects

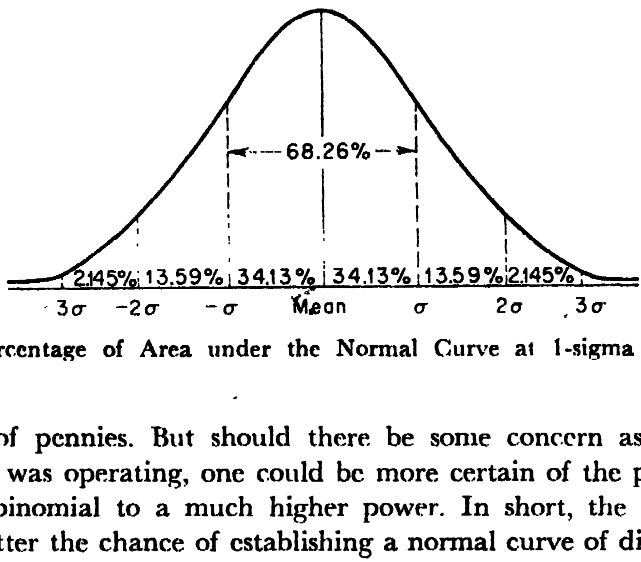


Fig. 18-2. Percentage of Area under the Normal Curve at 1-sigma Units on the Base Line.

the tossing of pennies. But should there be some concern as to whether pure chance was operating, one could be more certain of the prediction by raising the binomial to a much higher power. In short, the more repetitions, the better the chance of establishing a normal curve of distribution.

In the above curve all measures of central tendency are assumed to fall at the same place. Because the mean is the most accurate of the three measures, the standard deviation is calculated to show the spread of scores about the mean. One should form the habit of thinking of the standard deviation as a distance or length to be measured in either direction from the mean. This distance is laid out under the curve as shown in Figure 18-2.

This shows that between a -1σ and a $+1\sigma$ there are 68.26 per cent of all cases under the curve. It also shows that between a -2σ and a $+2\sigma$ there are 95.44 per cent of the cases. Between a -3σ and a $+3\sigma$ lie over 99 per cent of all scores.

The formula for finding the standard deviation from *ungrouped scores* is

$$\sigma = \sqrt{\frac{\Sigma d^2}{N}}$$

Here, six relatively easy steps are followed:

1. Add the scores and divide by the number (N) in order to find the mean.
2. Find how far each score deviates from the mean.
3. Square each of these deviations.
4. Add each of the squares.
5. Divide by N.
6. Extract the square root.

To illustrate, the total raw scores on thirteen archery tests total 429. Dividing 429 by 13 gives a mean of 33.

Scores	d	d^2
39	6	36
38	5	25
37	4	16
36	3	12
35	2	4
34	1	1
31	2	4
31	2	4
31	2	4
29	4	16
28	5	25
27	6	36
429		183

Proceeding from here, it is observed that 39 deviates 6 from 33, 38 deviates 5 from 33, 37 deviates 4 from 33, etc. Each deviation is then squared in order to obtain the values in the d^2 column. The sum of the d^2 column is 183. This figure divided by $N = 183/13$, or 14. Taking the square root of 14, a standard deviation of 3.73 is obtained.

If this standard deviation of 3.73 is subtracted from the mean, 33, and also added to the mean, two figures are obtained, 29.27 and 36.73, respectively. This shows that approximately two-thirds of the cases lie within the limits of 29.27 and 36.73 score points on either side of the mean.

The formula for finding the standard deviation with *grouped data* is set up in Table 18-4.

Observe that Table 18-4 is the same as Table 18-3, except for the new column, fd^2. The Σfd has already been calculated as -9. The mean has already been determined as 51.1.

In arriving at the standard deviation as in Table 18-4, these steps should be followed:

1. Carry out the same steps as were followed in calculating the mean.
2. Add another column (fd^2), which represents the "d" column times

the "fd" column. Actually, this represents the particular frequency times its squared deviation.

3. Add the "fd^2" column. All numbers are now positive.
4. Compute $(\Sigma fd/N)^2$ and subtract this figure from $\Sigma fd^2/N$.
5. Substitute the figures obtained in the formula and obtain the square root.
6. Multiply the square root figure by the size of the step interval (5). The product represents the standard deviation.

Table 18-4
Calculation of the Standard Deviation*

Scores	Mid-point	f	d	fd	fd^2
89.5–94.4	92	1	8	8	64
84.5–89.4	87	2	7	14	98
79.5–84.4	82	3	6	18	108
74.5–79.4	77	3	5	15	75
69.5–74.4	72	2	4	8	32
64.5–69.4	67	2	3	6	18
59.5–64.4	62	3	2	6	12
54.5–59.4	57	4	1	4	4
49.5–54.4	52	4	0	0	0
44.5–49.4	47	7	−1	−7	7
39.5–44.4	42	4	−2	−8	16
34.5–39.4	42	4	−3	−12	36
29.5–34.4	32	3	−4	−12	48
24.5–29.4	27	3	−5	−15	75
19.5–24.4	22	2	−6	−12	72
14.5–19.4	17	2	−7	−14	98
9.5–14.4	12	1	−8	−8	64
					727 = Σfd^2

* $N = 50$.

$$\sigma = i\sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2}$$
$$= 5\sqrt{\frac{727}{50} - \left(\frac{-9}{50}\right)^2}$$
$$= 5\sqrt{14.54 - .0324}$$
$$= 5\sqrt{14.50}$$
$$= 5 \times 3.80$$
$$= 19$$

With this it is possible to tell the limit of variability for the scores and students possessing the scores. Using the standard deviation, several sets of data from different experiments or tests can be compared. Test results with about the same mean and standard deviation are very much alike in con-

centration and variability of scores. With a standard deviation, groups and individuals may be properly equated. In Table 18-4, the standard deviation of 19 shows that between a -1σ and a $+1\sigma$ there are thirty-eight scores (68.26 per cent). Since the mean is 51.1, these thirty-eight scores will fall 19 scores above and 19 scores below the mean. The standard deviation, therefore, would be from score 32 to 70.

The Standard Score

Very often it is necessary to interpret a score in terms of its position in the distribution. If we know the particular score, the mean, and the standard deviation, we can state the score in terms of its exact position in the distribution. This becomes the standard score or *Z score*. To determine the Z score, find the distance of a particular score (X) from the mean (M) and divide this by the standard deviation (σ). For example, in Table 18-4 the mean is 51.5, and the standard deviation is 19. Suppose we wanted to know where in the distribution a score of 80 would fall. We would proceed as follows:

$$\text{Z score} = \frac{X - M}{\sigma} = \frac{80 - 51.5}{19} = 1.5$$

In this case the Z score of 1.5 shows that a score of 80 is halfway between one and two deviations above the mean.

Studies of variability between persons often use standard scores. A basis is provided for comparing scores. For example, the pupil with a raw score of 80 can be compared with a pupil with a raw score of 60. The question of how significant the difference of 20 points is in the scores can be answered in terms of the Z scores, which are 1.5 and .5, respectively. Thus the pupil with a raw score of 80 and Z score of 1.5 is halfway between one and two standard deviations above the mean with approximately 90 per cent of the class below him (see percentage area under the normal curve in Figure 18-2). The pupil with only 60 points and a Z score of .5 is only half of one standard deviation above the mean, or approximately at the sixty-seventh percentile. Thus the 20-point difference in the two pupil scores is quite significant, particularly for grading.

T Scores

Sometimes it is convenient to convert a score into a *T score*. In the T score the mean and middle of the distribution are always 50, a figure that is regularly understood because it is at the mid-point between 0 and 100.

$$\text{T score} = 50 + 10Z$$

In this case a Z score of 1.5 would convert into a T score of 65. There are many T-score tables accompanying physical education tests. The T score has an advantage over the Z score because the negative sign is eliminated

and the large unit of one standard deviation is broken down into units one-tenth as large. Thus an ordinary distribution of scores with a range of 5 to 65 on its base line yields 50 to 60 integral T-scale scores. Furthermore, the T score extends its scale over a spread of ten standard deviations, or 100 units in all. This is helpful when all groups from the lowest to the highest are considered on a single common scale. The 100-unit (or 10σ) scale is extensive enough to take care of any group of scores.

Building a Norm Table

In order to give meaning to the raw scores obtained on local health and physical education tests, it is necessary to group the scores, determine the central tendency and variability of the distribution, and arrange the scores in some useful order from the lowest to the highest.

What is needed is a common basis for comparing scores so pupils can be classified and related to one another in terms of achievement on written tests and performance measures.

Reference has already been made to norm tables based on percentiles or deciles. It is not difficult to build a percentile table from a set of raw data, and it is easy to interpret. However, the percentile table does not have equal scale units of the same length, and it does not reflect the variability of the distribution of test scores.

Two scales that have wide use in the building of norm tables are the sigma scale and the T scale.

THE SIGMA SCALE

This scale consists solely of three standard deviations on either side of the mean and is frequently referred to as the 6-sigma scale. To construct the sigma scale, proceed as follows:

1. Calculate the mean and standard deviation of the distribution of test scores.

2. On a piece of paper arrange the numbers 1 to 100 in a column. Write the mean of the scores next to number 50.

3. Multiply the standard deviation by 3 and divide by 50.

4. Add this number to the mean for calculating scale points 51 to 100 on the scale. Do this consecutively for each number. Likewise, consecutively subtract this number from the mean for establishing the values for each number from 49 to 0.

For example, suppose that in a distribution of 300 health test scores, ranging from 14 to 98 points, the mean score was 56 and the standard deviation 6 points. By following the instructions and multiplying the standard deviation by 3 and dividing by 50, we would obtain

$$\frac{6 \times 3}{50} = .36$$

By applying this to the scale, where the mean is 56 (at scale score 50), the scale score 51 would be 56 + .36, or 56.36. Likewise, scale score 52 would be 56.36 + .36. On the minus side, scale score 49 would be 56 — .36, or 55.64. Score 48 would then be 55.64 — .36, or 55.28.

The scale can be reduced from hundredths to tenths in instances where close differences are not needed. In such a case the scale unit figure, .36 in the above example, would become 3.6 (see Table 18-5).

Table 18-5
Illustration of Sigma-scale and T-scale Tables

Sigma scale			T scale	
Scale score	Test score	Rounded off	Scale score	Test score
100	74.0	74	100	86
90	70.4	70	90	80
80	66.8	67	80	74
70	63.2	63	70	68
60	59.6	60	60	62
50	56	56	50	56
40	52.4	52	40	50
30	48.8	49	30	44
20	45.2	45	20	38
10	41.6	42	10	32
0	38.0	38	0	26

Calculations

Mean (plus and minus): $\frac{3\sigma}{50}$

Hundredths: $\frac{3 \times 6}{50} = .36$

Tenths: $\frac{3 \times 6}{5} = 3.6$

Mean = 56

Calculations

Mean (plus and minus): $\frac{5\sigma}{50}$

Hundredths: $\frac{5 \times 6}{50} = .6$

Tenths: $\frac{5 \times 6}{5} = 6.0$

Mean = 56

THE T SCALE

The primary difference between the T scale and the sigma scale is that there are five standard deviations instead of three on either side of the mean. The same procedure employed to build the sigma scale is used, except that the standard deviation is multiplied by 5 and divided by 50. The same example that was given to illustrate the sigma scale is shown in Table 18-5 applied to the construction of the T-scale table.

Other arrangements for scaling the distribution of scores are available. The *Hull Scale,* for example, differs only slightly from the sigma and T scales. Here the standard deviation is multiplied by 3.5.

Reliability of Various Measures

It is common practice to expect errors to appear in any process of measurement and evaluation. It has already been pointed out that reliability is dependent on the procedure followed by the person doing the testing, by the nature of the measuring instruments, and by the behavior of the person being tested. In addition, there are certain errors that are peculiar to the statistical processes used to evaluate test results. It is practically impossible, therefore, to construct a test with perfect reliability. This is especially true in health and physical education, where many tests and measuring devices are not refined enough to escape the effect of numerous variables.

In experimental research, where a control group is being related to an experimental group, it is often desirable to have both groups as near equal as possible in characteristics and test scores. Yet the statistics that are used to help equate the groups in central tendency and variability are subject to error. For example, an obtained mean for one group of scores has a certain error, which is called the *standard error of the mean.* This means simply that the sample of scores used to calculate the mean is a fallible sample; it is subject to error.[5] The formula for the standard error of the mean is as follows:

$$\sigma_M = \frac{\sigma}{\sqrt{N}}$$

where σ_M = standard error of the mean
σ = standard deviation
M = mean
N = number of cases

Usually where there are more cases (N), the σ_M is lower and there will be greater accuracy of the obtained scores. To illustrate this, suppose that in a group of scores the mean was 50, the standard deviation 8, and the number of cases 64. Here the σ_M would be exactly ($8/\sqrt{64} = 1$). This indicates that the mean is reliable to the extent that it will not vary more than 1 point on either side of the mean in at least 68 per cent of the cases (one standard deviation). The true mean, therefore, will fall between 49 and 51. Suppose, however, that instead of 64 cases there had been 100 cases. The increased number of cases lowers the standard error of the mean:

$$\frac{8}{\sqrt{100}} = .8$$

In this example, reliability has been increased, for the true mean will now fall between .8 points of either side of the mean (49.2–50.8).

[5] A good sampling of subjects or scores (which has quite a bearing on reliability) is brought about by choosing subjects or scores at random from a large population or group so they are characteristic of the whole population.

When comparing two groups it is necessary to find the σ_M of each group. With these figures it is possible to find the difference between the σ_M for each group. By examining the σ_M for each group the actual difference (d) can readily be seen. But this figure is not quite correct. It is necessary to calculate the *standard error of the difference.* The formula for the standard error of a difference is as follows:

$$\sigma_d = \sqrt{\sigma_{M_1}^2 + \sigma_{M_2}^2}$$

where σ_d = standard error of difference between two means
$\sigma_M{}^2$ = standard error of the mean squared
1, 2 = subscripts to distinguish between the two groups

When the standard error of the difference (σ_d) is obtained, it tells the research worker the true difference between the two groups of scores. The next question is whether this difference is a significant one. If it is significant, the two groups probably should not be compared with each other for they are not very much alike. To ascertain whether the standard error of the difference is significant the σ_d is divided into the actual difference (d).

$$\frac{d}{\sigma_d}$$

When the d/σ_d is greater than 3 (3 sigma), the obtained difference is reliable. This is because practically all cases in a normal distribution lie within 3 sigma (three standard deviations) on each side of the mean.

The Coefficient of Reliability

As already indicated in Chapter 3, test reliability is extremely important. The makers of standardized tests ordinarily report reliability by calculating the *coefficient of correlation,* which is known in this instance as the *coefficient of reliability.* Usually a sample of 200 cases will be sufficient for calculating the coefficient of reliability.[6]

There are a number of ways to obtain the coefficient Where a test has two equivalent forms, one form may be given one day and a day or two later the alternate form may be given. The correlation between test scores on each form determines the reliability coefficient.

Frequently, when a test has only one form it is given on two occasions. With knowledge tests this will not always work; too much may be remembered between tests. But in performance tests, such as sit-ups, pull-ups, or manuometer measures, this practice is satisfactory, provided fatigue or disinterest is not a factor.

Another way to check reliability is to divide the test into two equal parts. For example, the odd-numbered items can be correlated with the even-numbered items for each student and the reliability coefficient between the

[6] M. J. Nelson and E. C. Denny, *Statistics for Teachers,* New York, The Dryden Press, Inc., 1940, p. 140.

two parts obtained. Here the resulting coefficient is considered reliable only for a test that is half as long as the original test. Thus to determine the reliability for the whole test, the coefficient must be boosted. This is accomplished through the use of the Spearman-Brown prophecy formula, which is

$$r_x = \frac{nr}{1 + (n - 1)r}$$

where r_x = coefficient sought
n = number of times whole test is longer than the parts
r = coefficient of correlation between parts

For example, to calculate the reliability of a test for which the correlation between odds and evens is $r = .60$, substitute in the formula as follows:

$$r_x = \frac{2 \times .60}{1 + .60} = \frac{1.20}{1.60} = .75$$

Thus the Spearman-Brown prophecy formula has shown that by doubling the length of the test the coefficient of reliability can be increased from .60 to .75.

Correlation

To be effective in evaluating test scores or behavior traits, it is often quite important to have a knowledge of the relationship that exists between one capacity and another. A health teacher might want to know whether or not there is any relationship between scores obtained on a nutrition knowledge test and scores obtained on a test of eating habits. A physical education instructor might want to know whether there is a relationship between long arms and the ability to throw the discus. Many statistical methods exist for finding relationship or correlation.

The designation for the coefficient of correlation is the letter r, which may vary all the way from a $+1$ to -1; that is, from a perfect positive relationship to a perfect negative relationship. If, for example, it was found that tall persons are usually better high jumpers than short persons, one would say that there is some relationship between height and high-jumping ability, which would be expressed as a coefficient of correlation. In this example the r would be somewhere between 0.0 and $+1.0$. If the coefficient of correlation was zero, it would indicate a complete lack of relationship. This is not a percentage but simply an index figure.

The *Spearman rank method of correlation* is convenient for many types of data. It is a means to rank relatively two sets of test scores. The degree of correlation is dependent upon the degree of agreement between the two rankings. When the ranking is completed, the Spearman formula is applied to obtain the correlation. Here the Greek rho (ρ) signifies the coefficient of correlation.

$$\rho = 1 - \frac{6\Sigma D^2}{N(N^2 - 1)}$$

Where a large number of cases or scores is involved, the rank method is not only cumbersome but also inaccurate. By the use of the ***Pearson product-moment*** method a large number of test cases can be compared with another group to ascertain whether a positive or negative relationship exists between the two. Also where there are many cases there will undoubtedly be a better sampling of scores, and any relationship that is found will be more meaningful.

Study the correlation chart (Table 18-6) to see how it is ruled. Notice that the scattergram is laid off with respect to the X axis (horizontal) and the Y axis (vertical), which cross each other near the center. Here, the nutrition knowledge scores (Y) are being contrasted with the eating habit scores (X).

All fifty scores for each test have been tallied into the proper cell for each pair of performances. The tallies have been placed in the upper left-hand corner. Then a guessed mean was selected approximately in the center of each distribution of scores, and the cells in the row and column were ruled off with heavy pencil ($X = 150$ to 159; $Y = 50$ to 59). By adding the rows of tally marks, the "fy" and "fx" columns are filled. These are the frequencies for the Y and X variables, which, of course, add up to the fifty scores. Next, the deviations (d) from the guessed mean are figured; they are then multiplied by the frequencies (fd) and squared (fd^2), and the columns are totaled for each variable. The correction (c) for each variable (C_x and C_y) was then computed ($C_x = \Sigma fd/N = .32$ and $C_y = \Sigma fd/N = .10$).

Notice that many cells contain a circled figure with a plus or minus sign. The figure indicates the size of the product-moment. It represents the number of cells away from the intersecting guessed mean lines multiplied by the number of tallies in the cell. Thus, the circled figure 8 in the cell (with two tallies) under X 170–179 and Y 70–79 is four cells away from the intersecting lines. Multiplying 4 by 2 tallies gives the weight of 8. The circled figures in the four areas in point of direction from the intersecting guessed mean lines are always positive in the right area above and left area below. They are always negative in the left area above and right area below.

The product-moment of all cells was computed by adding both the plus and minus circled figures and entering the result in the appropriate plus or minus box under Σxy. The algebraic total of the Σxy was then found

$$\Sigma xy = 158$$

Notice that by calculating Σxy under *both* the X and Y variables it is possible to check for accuracy. The Σxy, 158, appears the same under both variables. At this point the standard deviation for each distribution was calculated, and the formula for r was applied. By substituting scattergram information in the formula, an r of .74 was found.

This correlation, being positive, would indicate that there is some degree of correlation between nutritional knowledge and eating habits. Conclusions,

Table 18-6
Calculation of the Product-moment Coefficient of Correlation between Scores on a Nutrition Knowledge Test and a Test on Eating Habits

X-variable: eating habits

Y-variable: nutrition knowledge

	90-99	100-109	110-119	120-129	130-139	140-149	150-159	160-169	170-179	180-189	190-199	*fy*	*d*	*fd*	*fd*²	Σ*xy* +	Σ*xy* −
90-99																	
80-89									/ 6	// 18	// 24	5	+3	+15	45	48	
70-79						/ −2	/	/// 6	// 8	/ 6	/ 8	9	+2	+18	36	28	2
60-69						/ −1	//	/// 3	// 4	/ 3		9	+1	+9	9	10	1
50-59				/	/		////		///	/		10					
40-49				/ 3	/// 6	/// 3	/	// −2				10	−1	−10	10	12	2
30-39			/// 24		// 8							5	−2	−10	20	32	
20-29	/ 18	/ 15										2	−3	−6	18	33	
fx	1	1	3	2	6	5	8	8	8	5	3	50		+16	138	+163	−5
d	−6	−5	−4	−3	−2	−1		+1	+2	+3	+4	.				Σ*xy* = +158	
fd	−6	−5	−12	−6	−12	−5		+8	+16	+15	+12	+5					
*fd*²	36	25	48	18	24	5		8	32	45	48	289					
Σ*xy* +	18	15	24	3	14	3		9	18	27	32	163	Σ*xy* = +158				
Σ*xy* −						3		2				−5					

$$r = \frac{\frac{\Sigma xy}{N} - (C_x C_y)}{(\sigma_x)(\sigma_y)}$$

$$= \frac{\frac{158}{50} - (.32)(.1)}{(1.6)(2.4)}$$

$$= \frac{3.16 - .31}{3.84}$$

$$= \frac{2.85}{3.84}$$

$$= .74$$

where r = coefficient of correlation

xy = algebraic sum of the deviation from the guessed mean in terms of the step interval

N = 50 cases

C_x = corrections in X: $\Sigma fd/N = +16/50 = .32$

C_y = corrections in Y: $\Sigma fd/N = +5/50 = .10$

σ_x = standard deviation of X scores in terms of the step interval

$$= \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2}$$

$$= \sqrt{\frac{138}{50} - .1024}$$

$$= 1.6$$

σ_y = standard deviation of Y scores in terms of the step interval

$$= \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2}$$

$$= \sqrt{\frac{289}{50} - .01}$$

$$= 2.4$$

however, from data of fifty cases should not be given great weight. Several hundred cases would afford a better sampling.

Before leaving the topic of correlation it should be added that several other types of correlation are useful in health and physical education, especially in research. The more advanced correlation methods will be found in a basic statistics textbook.

SELECTED REFERENCES

Garrett, H. E.: *Statistics in Psychology and Education*, 5th ed., New York, Longmans, Green & Co., Inc., 1958.

Good, Carter V.: *Introduction to Education Research*, New York, Appleton-Century-Crofts, Inc., 1959.

Guilford, J. P.: *Fundamental Statistics in Psychology and Education*, 3d ed., New York, McGraw-Hill Book Company, Inc., 1956.

Ingraham, Hollis S.: "Statistics and Medical Knowledge," *American Journal of Public Health*, 48:1449–1459, November 1958.

Monroe, Walter S., and Max D. Englehardt: *The Scientific Study of Educational Problems*, New York, The Macmillan Company, 1936.

Appendix A. Scoring Tables

Table A-1
AAHPER Youth Fitness Test
Percentile Scores for College Men

Percentile	Sit-ups	Pull-ups	St. broad jump	Shuttle	50-yd dash	Softball throw	600-yd run
100	100	20	9-6	8.3	5.5	315	1:12
95	99	12	8-5	9.0	6.1	239	1:35
90	97	10	8-2	9.1	6.2	226	1:38
85	79	10	7-11	9.1	6.3	217	1:40
80	68	9	7-10	9.2	6.4	211	1:42
75	61	8	7-8	9.4	6.5	206	1:44
70	58	8	7-7	9.5	6.5	200	1:45
65	52	7	7-6	9.5	6.6	196	1:47
60	51	7	7-5	9.6	6.6	192	1:49
55	50	6	7-4	9.6	6.7	188	1:50
50	47	6	7-3	9.7	6.8	184	1:52
45	44	5	7-1	9.8	5.8	180	1:53
40	41	5	7-0	9.9	6.9	176	1:55
35	38	4	6-11	10.0	7.0	171	1:57
30	36	4	6-10	10.0	7.0	166	1:59
25	34	3	6-9	10.1	7.1	161	2:01
20	31	3	6-7	10.2	7.1	156	2:05
15	29	2	6-5	10.4	7.1	150	2:09
10	26	1	6-2	10.6	7.5	140	2:15
5	22	0	5-10	11.1	7.7	125	2:25
0	0	0	4-2	13.9	9.1	55	3:43

SOURCE: American Association for Health, Physical Education, and Recreation, *AAHPER Youth Fitness Test,* Washington, 1960.

Table A-2
Badminton Achievement Scales (French-Stalter)

Rating	Serve	Clear	Wrist volley	Smash	Shuttle
Superior	57 and up	88 and up	98 and up	89 and up	38 and up
Good	42–56	70–87	77–97	75–88	35–37
Average	27–41	51–69	55–76	60–74	31–34
Poor	12–26	33–50	34–54	45–59	28–30
Inferior	11 and below	32 and below	33 and below	44 and below	27 and below

SOURCE: Esther French and Evelyn Stalter, "Study of Skill Tests in Badminton for College Women," *The Research Quarterly,* 20:257–272, October, 1949.

Table A-3
California Physical Performance Test Standards
Elementary School Boys and Girls (Grades Five to Eight)

Rating by class		Boys						Girls		
		Standing broad jump	Jump and reach	Pull-up	Push-up	50-yd dash	Throw for distance (softball)	Jump and reach	50-yd dash	Throw for distance (softball)
A	Superior	6–6	15	11	23	7.7	107	13	7.5	70
	Excellent	5–10	13	7	15	8.3	93	12	8.1	56
	Good	5–5	11	5	10	8.6	83	10	8.4	47
	Satisfactory	4–11	9	3	8	9.0	73	9	8.8	37
	Fair	4–3	6	2	4	9.5	59	7	9.4	23
	Poor	4–2	5	1	3	9.6	58	6	9.5	22
B	Superior	6–8	16	11	24	7.4	120	14	7.3	78
	Excellent	6–0	13	7	16	7.9	105	12	7.9	63
	Good	5–7	11	5	10	8.3	95	11	8.2	54
	Satisfactory	5–1	9	3	8	8.7	85	9	8.6	44
	Fair	4–5	7	2	4	9.2	70	7	9.2	29
	Poor	4–4	6	1	3	9.3	69	6	9.3	28
C	Superior	6–10	16	11	25	7.1	130	14	7.1	86
	Excellent	6–2	14	7	16	7.6	115	13	7.7	71
	Good	5–9	12	5	11	8.0	104	11	8.1	61
	Satisfactory	5–3	10	3	8	8.4	93	9	8.5	50
	Fair	4–7	7	2	5	8.9	77	7	9.0	35
	Poor	4–6	6	1	4	9.0	76	6	9.1	34
D	Superior	7–0	17	11	26	6.9	140	15	7.0	93
	Excellent	6–4	14	7	18	7.4	123	13	7.6	78
	Good	5–11	12	5	12	7.8	112	12	7.9	68
	Satisfactory	5–5	11	4	9	8.2	101	10	8.3	57
	Fair	4–9	8	2	5	8.7	84	8	8.9	42
	Poor	4–8	7	1	4	8.8	83	7	9.0	41
E	Superior	7–3	18	11	29	6.7	151	16	6.9	101
	Excellent	6–7	15	8	20	7.2	133	14	7.5	85
	Good	6–1	13	5	14	7.6	121	12	7.9	74
	Satisfactory	5–7	11	4	11	8.0	108	11	8.3	63
	Fair	4–11	8	2	6	8.5	90	8	8.8	47
	Poor	4–10	7	1	5	8.6	89	7	8.9	46
F	Superior	7–6	18	12	32	6.5	162	17	6.9	105
	Excellent	6–10	16	8	23	7.1	142	15	7.5	89
	Good	6–4	14	5	16	7.4	129	13	7.9	77
	Satisfactory	5–10	12	4	13	7.9	115	11	8.3	66
	Fair	5–2	9	2	8	8.4	96	9	8.8	49
	Poor	5–1	8	1	7	8.5	95	8	8.9	48
G	Superior	7–8	19	12	34	6.4	175	18	7.0	108
	Excellent	7–0	16	8	24	6.9	153	15	7.6	91
	Good	6–6	15	6	17	7.3	138	13	8.0	79
	Satisfactory	6–0	13	4	14	7.7	123	12	8.4	67
	Fair	5–4	10	2	9	8.2	101	9	8.9	50
	Poor	5–3	9	1	8	8.3	100	8	9.0	49

Table A-3 (*Continued*)
Junior, Senior, and Four-year High School Boys

Rating by class		Standing broad jump	Jump and reach	Pull-up	Push-up	50-yd dash	75-yd dash	Throw for distance (softball)
A	Superior	9–0	21	18	36	6.2	8.5	197
	Excellent	8–2	18	13	27	6.7	9.3	173
	Good	7–7	16	10	20	7.1	9.8	157
	Satisfactory	7–1	14	8	17	7.4	10.3	140
	Fair	6–3	11	5	13	8.0	11.1	116
	Poor	6–2	10	4	12	8.1	11.2	115
B	Superior	8–3	20	16	34	6.3	9.0	183
	Excellent	7–6	17	12	25	6.9	9.8	159
	Good	7–0	15	9	18	7.2	10.3	142
	Satisfactory	6–6	13	7	15	7.6	10.8	126
	Fair	5–9	10	4	11	8.1	11.6	102
	Poor	5–8	9	3	10	8.2	11.7	101
C	Superior	7–10	19	15	32	6.4	9.4	172
	Excellent	7–2	16	11	23	7.0	10.2	148
	Good	6–8	14	8	17	7.3	10.7	131
	Satisfactory	6–3	12	6	14	7.7	11.2	115
	Fair	5–7	9	3	9	8.2	12.0	91
	Poor	5–6	8	2	8	8.3	12.1	90
D	Superior	7–7	18	14	31	6.5	9.7	163
	Excellent	6–11	15	9	22	7.1	10.5	139
	Good	6–6	14	7	16	7.4	11.0	123
	Satisfactory	6–0	12	5	12	7.8	11.5	106
	Fair	5–4	9	2	8	8.3	12.3	82
	Poor	5–3	8	1	7	8.4	12.4	81
E	Superior	7–6	18	13	31	6.6	9.9	155
	Excellent	6–10	15	8	21	7.1	10.7	131
	Good	6–4	13	6	15	7.5	11.2	115
	Satisfactory	5–11	11	4	11	7.9	11.7	98
	Fair	5–2	8	2	7	8.4	12.5	74
	Poor	5–1	7	1	6	8.5	12.6	73
F	Superior	7–4	17	11	30	6.7	10.1	147
	Excellent	6–8	14	7	20	7.2	10.9	123
	Good	6–3	13	4	14	7.6	11.4	107
	Satisfactory	5–9	11	2	11	7.9	11.9	90
	Fair	5–1	8	1	7	8.5	12.7	66
	Poor	5–0	7	0	6	8.6	12.8	65

Table A-3 (*Continued*)
Junior, Senior, and Four-year High School Girls

Rating	Standing broad jump	Jump and reach	Knee push-up	Sit-up	50-yd dash	Shuttle	Throw for distance (softball)
Superior	6–10	16	45	48	7.1	9.5	118
Excellent	6–2	14	32	36	7.8	10.4	96
Good	5–9	12	25	29	8.2	11.0	80
Satisfactory	5–4	11	16	21	8.9	11.8	65
Fair	4–8	8	4	9	9.7	12.5	42
Poor	4 7	7	3	8	9.8	12.6	41

SOURCE: California State Department of Education, *California Physical Performance Test,* Bureau of Health Education, Physical Education, and Recreation, Sacramento, Calif., 1958.

Table A-4
Carpenter's Motor Ability Test Scores for First Three Grades (Boys)

	0	1	2	3	4	5	6	7	8	9
600	115.94	116.24	116.54	116.84	117.14	117.44	117.75	118.05	118.35	118.65
590	112.93	113.23	113.53	113.83	114.13	114.44	114.74	115.04	115.34	115.64
580	109.92	110.22	110.52	110.82	111.13	111.43	111.73	112.03	112.33	112.63
570	106.91	107.21	107.51	107.82	108.12	108.42	108.72	109.02	109.32	109.62
560	103.90	104.20	104.51	104.81	105.11	105.41	105.71	106.01	106.31	106.61
550	100.90	101.20	101.50	101.80	102.10	102.40	102.70	103.00	103.30	103.60
540	97.80	98.19	98.49	98.79	99.09	99.39	99.69	99.99	100.29	100.59
530	94.88	95.18	95.48	95.78	96.08	96.38	96.68	96.98	97.28	97.59
520	91.87	92.17	92.45	92.77	93.07	93.37	93.67	93.97	94.28	94.58
510	88.86	89.16	89.46	89.76	90.06	90.36	90.66	90.97	91.27	91.57
500	85.85	86.15	86.45	86.75	87.05	87.35	87.66	87.96	88.26	88.56
490	82.84	83.14	83.44	83.74	84.04	84.35	84.65	84.95	85.25	85.55
480	79.83	80.13	80.43	80.73	81.04	81.34	81.64	81.94	82.24	82.54
470	76.82	77.12	77.42	77.73	78.03	78.33	78.63	78.93	79.23	79.53
460	73.81	74.11	74.42	74.72	75.02	75.32	75.62	75.92	76.22	76.52
450	70.81	71.11	71.41	71.71	72.01	72.31	72.61	72.91	73.21	73.51
440	76.80	68.10	68.40	68.70	69.00	69.30	69.60	69.90	70.20	70.50
430	64.79	65.09	65.39	65.69	65.99	66.29	66.59	66.89	67.19	67.50
420	61.78	62.08	62.38	62.68	62.98	63.28	63.58	63.88	64.19	64.49
410	58.77	59.07	59.37	59.67	59.97	60.27	60.57	60.88	61.18	61.48
400	55.76	56.06	56.36	56.66	56.96	57.26	57.57	57.87	58.17	58.47

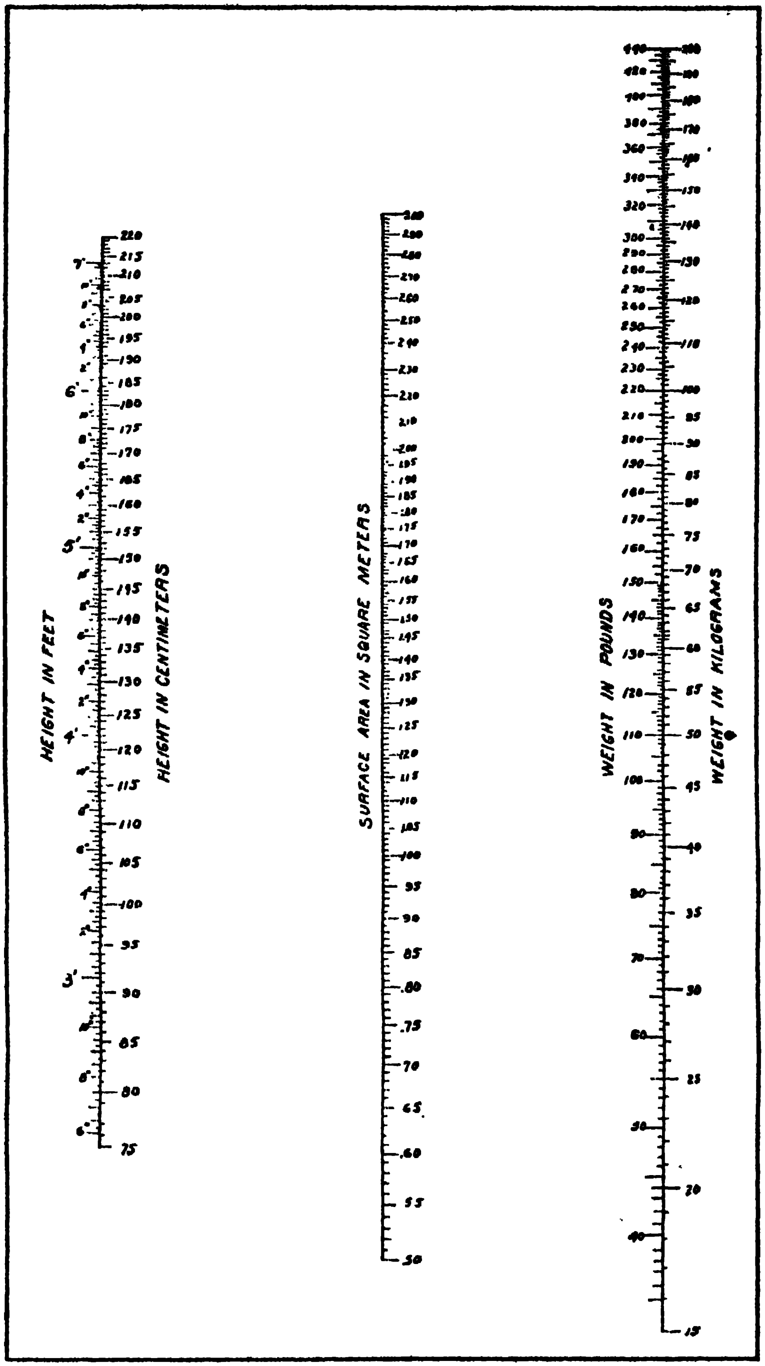
SOURCE: Aileen Carpenter, "Strength Testing in the First Three Grades," *The Research Quarterly,* 13:332–335, October, 1942.

Table A-5
Carpenter's Motor Ability Test Scores for First Three Grades (Girls)

	0	1	2	3	4	5	6	7	8	9
600	125.03	125.28	125.54	125.79	126.05	126.30	126.56	126.81	127.07	127.32
590	122.48	122.74	122.99	123.23	123.50	123.76	124.01	124.26	124.52	124.78
580	119.93	120.19	120.44	120.70	120.95	121.21	121.46	121.72	121.97	122.23
570	117.38	117.64	117.89	118.15	118.40	118.66	118.91	119.17	119.42	119.68
560	114.83	115.09	115.34	115.60	115.85	116.11	116.36	116.62	116.87	117.13
550	112.29	112.54	112.79	113.05	113.30	113.56	113.81	114.07	114.32	114.58
540	109.74	109.39	110.25	110.50	110.76	111.01	111 27	111.52	111.78	112.03
530	107.19	107.44	107.70	107.95	108.21	108.46	108.72	108.97	109.32	109.48
520	104.64	104.89	105.15	105.40	105.66	105.91	106.17	106.42	106.68	106.93
510	102.09	102.34	102.60	102.85	103.11	103.36	103.62	103.87	104.13	104.38
500	99.54	99.79	100.95	100.30	100.55	100.81	101.07	101.32	101.58	101.83
490	96.99	97.25	97.50	97.76	98.01	98.27	98.52	98.78	99.03	99.29
480	94.44	94.70	94.95	95.21	95.46	95.72	95.97	96.23	96.48	96.74
470	91.89	92.15	92.40	92.66	92.91	93.17	93.42	93.68	93.93	94.19
460	89.34	89.60	89.85	90.11	90.36	90.62	90.87	91.13	91.38	91.64
450	86.80	87.05	87.30	87.56	87.81	88.07	88.32	88.58	88.83	89.09
440	84.25	84.50	84.76	85.01	85.27	85.52	85.78	86.03	86.29	86.54
430	81.70	81.95	82.21	82.46	82.72	82.97	83.23	83.48	83.74	83.99
420	79.15	79.40	79.66	79.91	80.17	80.42	80.68	80.93	81.19	81.44
410	76.60	76.85	77.11	77.36	77.62	77.87	78.13	78.38	78.64	78.89
400	74.05	74.30	74.56	74.81	75.07	75.32	75.58	75.83	76.09	76.34

SOURCE: Aileen Carpenter, "Strength Testing in the First Three Grades," *The Research Quarterly,* 13:332–335, October, 1942.

Table A-6
Dubois Body Surface Chart (Nomograph) *



* The estimated body surface area is determined from the height and weight measurements located as points in the outside columns. A ruler placed so as to connect these points will intercept the center scale at the point indicating the equivalent surface area in square meters.

Table A-7
Dyer Sigma-scale Values for Test of Tennis Ability

Sigma scale	Test score	Sigma scale	Test score	Sigma scale	Test score	Sigma scale	Test score
100	67	75	50	50	33	25	16
99	66	74	49	49	32	24	15
98	...	73	...	48	...	23	
97	65	72	48	47	31	22	14
96	65	71	47	46	30	21	13
95	...	70	...	45	...	20	
94	63	69	46	44	29	19	12
93	62	68	45	43	28	18	11
92	...	67	44	42	27	17	10
91	61	66	...	41	...	16	
90	60	65	43	40	26	15	9
89	59	64	42	39	25	14	8
88	...	63	...	38	..	13	
87	58	62	41	37	24	12	7
86	57	61	40	36	23	11	6
85	...	60	...	35	...	10	
84	56	59	39	34	22	9	5
83	55	58	38	33	21	8	4
82	...	57	..	32	...	7	
81	54	56	37	31	20	6	3
80	53	55	36	30	19	5	2
79	...	54	...	29	...	4	
78	52	53	35	28	18	3	1
77	51	52	34	27	17	2	
76	...	51	..	26	...	1	

SOURCE: Joanna T. Dyer, "Revision of Backboard Test of Tennis Ability," *The Research Quarterly,* 9:25–31, March, 1939.

Table

Flanagan's Indices of

A Table of the Values of the Product-moment Coefficient of Correlation in a

Proportion of successes in the 27% scoring

Proportion of successes in the 27% scoring lowest on the continuous variable

	01	02	04	06	08	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48
01	0	11	23	30	35	40	43	46	49	51	53	55	57	59	61	62	63	65	66	67	68	69	70	71	72
02	-11	0	12	19	25	30	34	37	40	43	46	48	50	51	53	55	56	58	59	31	62	63	64	66	67
04	-23	-12	0	08	14	19	23	26	30	33	36	38	40	42	44	46	48	49	51	53	54	56	57	58	60
06	-30	-19	-08	0	06	11	15	19	23	26	29	31	33	36	38	40	42	44	45	47	48	50	52	53	55
08	-35	-25	-14	-06	0	05	09	13	17	20	23	25	28	30	32	35	37	38	40	42	44	45	47	49	51
10	-40	-30	-19	-11	-05	0	04	08	12	15	18	21	23	26	28	30	32	34	36	38	40	41	43	45	47
12	-43	-34	-23	-15	-09	-04	0	04	07	11	13	16	19	21	24	26	28	30	32	34	36	38	39	41	43
14	-46	-37	-26	-19	-13	-08	-04	0	03	07	10	12	15	18	20	22	25	27	29	31	33	34	36	38	40
16	-49	-40	-30	-23	-17	-12	-07	-03	0	03	06	09	12	14	17	19	21	24	26	28	30	31	33	35	37
18	-51	-43	-33	-26	-20	-15	-11	-07	-03	0	03	06	08	11	13	16	18	20	23	25	27	28	30	32	34
20	-53	-46	-36	-29	-23	-18	-13	-10	-06	-03	0	03	06	08	11	13	15	17	19	22	24	26	27	29	31
22	-55	-48	-38	-31	-25	-21	-16	-12	-09	-06	-03	0	03	06	08	10	12	15	17	19	21	23	25	27	29
24	-57	-50	-40	-33	-28	-23	-19	-15	-12	-08	-06	-03	0	03	05	08	10	12	14	16	18	20	22	24	26
26	-59	-51	-42	-36	-30	-26	-21	-18	-14	-11	-08	-06	-03	0	02	05	07	09	12	14	16	18	20	22	24
28	-61	-53	-44	-38	-32	-28	-24	-20	-17	-13	-11	-08	-05	-02	0	02	04	07	09	11	13	15	17	19	21
30	-62	-55	-46	-40	-35	-30	-26	-22	-19	-16	-13	-10	-08	-05	-02	0	02	04	07	09	11	13	15	17	19
32	-63	-56	-48	-42	-37	-32	-28	-25	-21	-18	-15	-12	-10	-07	-04	-02	0	02	04	07	09	11	13	15	17
34	-65	-58	-49	-44	-38	-34	-30	-27	-24	-20	-17	-15	-12	-09	-07	-04	-02	0	02	04	06	09	11	13	15
36	-66	-59	-51	-45	-40	-36	-32	-29	-26	-23	-19	-17	-14	-12	-09	-07	-04	-02	0	02	04	06	08	11	13
38	-67	-61	-53	-47	-42	-38	-34	-31	-28	-25	-22	-19	-16	-14	-11	-09	-07	-04	-02	0	02	04	06	08	11
40	-68	-62	-54	-48	-44	-40	-36	-33	-30	-27	-24	-21	-18	-16	-13	-11	-09	-06	-04	-02	0	02	04	06	08
42	-69	-63	-56	-50	-45	-41	-38	-34	-31	-28	-26	-23	-20	-18	-15	-13	-11	-09	-06	-04	-02	0	02	04	06
44	-70	-64	-57	-52	-47	-43	-39	-36	-33	-30	-27	-25	-22	-20	-17	-15	-13	-11	-08	-06	-04	-02	0	02	04
46	-71	-66	-58	-53	-49	-45	-41	-38	-35	-32	-29	-27	-24	-22	-19	-17	-15	-13	-11	-08	-06	-04	-02	0	02
48	-72	-67	-60	-55	-51	-47	-43	-40	-37	-34	-31	-29	-26	-24	-21	-19	-17	-15	-13	-11	-08	-06	-04	-02	0
50	-72	-68	-61	-56	-52	-48	-45	-42	-39	-36	-33	-31	-28	-26	-23	-21	-19	-17	-15	-13	-10	-08	-06	-04	-02
52	-73	-69	-62	-57	-53	-50	-46	-43	-40	-38	-35	-33	-30	-28	-26	-23	-21	-19	-17	-15	-12	-10	-08	-06	-04
54	-74	-70	-63	-59	-55	-51	-48	-45	-42	-39	-37	-34	-32	-30	-27	-25	-23	-21	-19	-16	-14	-12	-10	-08	-06
56	-75	-71	-64	-60	-56	-53	-49	-47	-44	-41	-39	-36	-34	-32	-29	-27	-25	-23	-21	-18	-16	-14	-12	-10	-08
58	-76	-72	-66	-61	-58	-54	-51	-48	-45	-43	-40	-38	-36	-33	-31	-29	-27	-25	-22	-20	-18	-16	-14	-12	-10
60	-77	-73	-67	-62	-59	-56	-52	-50	-47	-45	-42	-40	-37	-35	-33	-31	-29	-27	-25	-22	-21	-18	-16	-14	-12
62	-78	-73	-68	-64	-60	-57	-54	-51	-49	-47	-44	-42	-39	-37	-35	-33	-31	-29	-27	-25	-22	-20	-18	-16	-15
64	-78	-74	-69	-65	-61	-58	-55	-53	-50	-48	-46	-43	-41	-39	-37	-35	-33	-31	-29	-27	-25	-22	-21	-19	-17
66	-79	-75	-70	-66	-63	-60	-57	-54	-52	-49	-47	-45	-43	-41	-39	-37	-35	-33	-31	-29	-27	-25	-23	-21	-19
68	-80	-76	-71	-67	-64	-61	-58	-56	-53	-51	-49	-47	-45	-42	-40	-38	-37	-35	-33	-31	-29	-27	-25	-23	-21
70	-81	-77	-72	-68	-65	-63	-60	-57	-55	-53	-51	-49	-46	-44	-42	-40	-38	-37	-35	-33	-31	-29	-27	-25	-23
72	-82	-78	-73	-70	-66	-64	-61	-59	-57	-54	-52	-50	-48	-46	-44	-42	-40	-39	-37	-35	-33	-31	-29	-27	-26
74	-82	-79	-74	-71	-68	-65	-63	-60	-58	-56	-54	-52	-50	-48	-46	-44	-42	-41	-39	-37	-35	-33	-32	-30	-28
76	-83	-80	-75	-72	-69	-67	-64	-62	-60	-58	-56	-54	-52	-50	-48	-46	-45	-43	-41	-39	-37	-36	-34	-32	-30
78	-83	-80	-76	-73	-70	-68	-66	-63	-61	-60	-57	-56	-54	-52	-50	-49	-47	-45	-43	-42	-40	-38	-36	-34	-33
80	-84	-81	-77	-74	-72	-70	-67	-65	-63	-61	-60	-57	-56	-54	-52	-51	-49	-47	-46	-44	-42	-40	-39	-37	-35
82	-85	-82	-78	-76	-73	-71	-69	-67	-65	-63	-61	-60	-58	-56	-54	-53	-51	-49	-48	-47	-45	-43	-41	-39	-38
84	-86	-83	-80	-77	-75	-72	-70	-68	-67	-65	-63	-61	-60	-58	-57	-55	-53	-52	-50	-49	-47	-45	-44	-42	-40
86	87	-84	-81	-78	-76	-74	-72	-70	-68	-67	-65	-63	-62	-60	-59	-57	-56	-54	-53	-51	-50	-48	-47	-45	-43
88	-87	-85	-82	-80	-77	-76	-73	-72	-70	-69	-67	-66	-64	-63	-61	-60	-58	-57	-55	-54	-52	-51	-49	-48	-46
90	-88	-86	-83	-81	-79	-77	-76	-74	-72	-71	-70	-68	-67	-65	-64	-63	-61	-60	-58	-57	-56	-54	-53	-51	-50
92	-89	-87	-84	-82	-81	-79	-77	-76	-75	-73	-72	-70	-69	-68	-66	-65	-64	-63	-61	-60	-59	-58	-56	-55	-53
94	-90	-88	-86	-84	-82	-81	-80	-78	-77	-76	-74	-73	-72	-71	-70	-68	-67	-66	-65	-64	-62	-61	-60	-59	-57
96	-91	-90	-88	-86	-84	-83	-82	-81	-80	-78	-77	-76	-75	-74	-73	-72	-71	-70	-69	-68	-67	-66	-64	-63	-62
98	-92	-91	-90	-88	-87	-86	-85	-84	-83	-82	-81	-80	-80	-79	-78	-77	-76	-75	-74	-73	-73	-72	-71	-70	-69
99	-93	-92	-91	-90	-89	-88	-87	-87	-86	-85	-84	-83	-83	-82	-82	-81	-80	-79	-78	-78	-77	-76	-75	-74	-73

* Prepared by John C. Flanagan, January, 1936. Reproduced with his permission.

A-8
Discrimination:
Normal Bivariate Population Corresponding to Given Proportions of Successes*

highest on the continuous variable

50	52	54	56	58	60	62	64	66	68	70	72	74	76	78	80	82	84	86	88	90	92	94	96	98	9
72	73	74	75	76	77	78	78	79	80	81	82	82	83	83	84	85	86	87	86	88	89	90	91	92	9
68	69	70	71	72	73	73	74	75	76	77	78	79	80	80	81	82	83	84	85	86	87	88	90	91	9
61	62	63	64	66	67	68	69	70	71	72	73	74	75	76	77	78	80	81	82	83	84	86	88	90	9
56	57	59	60	61	62	64	65	66	67	68	70	71	72	73	74	76	77	78	80	81	82	84	86	88	90
52	53	55	56	58	59	60	61	63	64	65	66	68	69	70	72	73	75	76	77	79	81	82	84	87	89
48	50	51	53	54	56	57	58	60	61	63	64	65	67	68	70	71	72	74	76	77	79	81	83	86	88
45	46	48	49	51	52	54	55	57	58	60	61	63	64	66	67	69	70	72	73	76	77	80	82	85	87
42	43	45	47	48	50	51	53	54	56	57	59	60	62	63	65	67	68	70	72	74	76	78	81	84	87
39	40	42	44	45	47	49	50	52	53	55	57	58	60	61	63	65	67	68	70	72	75	77	80	83	86
36	37	39	41	43	45	47	48	49	51	53	54	56	58	60	61	63	65	67	69	71	73	76	78	82	85
33	35	37	39	40	42	44	46	47	49	51	52	54	56	57	60	61	63	65	67	70	72	74	77	81	84
31	33	34	36	38	40	42	43	45	47	49	50	52	54	56	57	60	61	63	66	68	70	73	76	80	8
28	30	32	34	36	37	39	41	43	45	46	48	50	52	54	56	58	60	62	64	67	69	72	75	80	83
26	28	30	32	33	35	37	39	41	42	44	46	48	50	52	54	56	58	60	63	65	68	71	74	79	82
23	26	27	29	31	33	35	37	39	40	42	44	46	48	50	52	54	57	59	61	64	66	70	73	78	82
21	23	25	27	29	31	33	35	37	38	40	42	44	46	49	51	53	55	57	60	63	65	68	72	77	81
19	21	23	25	27	29	31	33	35	37	38	40	42	45	47	49	51	53	56	58	61	64	67	71	76	80
17	19	21	23	25	27	29	31	33	35	37	39	41	43	45	47	49	52	54	57	60	63	66	70	75	79
15	17	19	21	22	25	27	29	31	33	35	37	39	41	43	46	48	50	53	55	58	61	65	69	74	78
13	15	16	18	20	22	25	27	29	31	33	35	37	39	42	44	47	49	51	54	57	60	64	68	73	78
10	12	14	16	18	21	22	25	27	29	31	33	35	37	40	42	45	47	50	52	56	59	62	67	73	77
08	10	12	14	16	18	20	22	25	27	29	31	33	36	38	40	43	45	48	51	54	58	61	66	72	76
06	08	10	12	14	16	18	21	23	25	27	29	32	34	36	39	41	44	47	49	53	56	60	64	71	75
04	06	08	10	12	14	16	19	21	23	25	27	30	32	34	37	39	42	45	48	51	55	59	63	70	74
02	04	06	08	10	12	15	17	19	21	23	26	28	30	33	35	38	40	43	46	50	53	57	62	69	73
0	02	04	06	08	10	13	15	17	19	21	23	26	28	31	33	36	39	42	45	48	52	56	61	68	72
-02	0	02	04	06	08	11	13	15	17	19	21	24	26	29	31	34	37	40	43	47	51	55	60	67	72
-04	-02	0	02	04	06	08	11	13	15	17	19	22	24	27	29	32	35	38	41	45	49	53	58	66	71
-06	-04	-02	0	02	04	06	08	11	13	15	17	20	22	25	27	30	33	36	39	43	47	52	57	64	70
-08	-06	-04	-02	0	02	04	06	09	11	13	15	18	20	23	26	28	31	34	38	41	45	50	56	63	69
-10	-08	-06	-04	-02	0	02	04	06	09	11	13	16	18	21	24	27	30	33	36	40	44	48	54	62	68
-13	-11	-08	-06	-04	-02	0	02	04	07	09	11	14	16	19	22	25	28	31	34	38	42	47	53	61	67
-15	-13	-11	-08	-06	-04	-02	0	02	04	07	09	12	14	17	19	23	26	29	32	36	40	45	51	59	66
-17	-15	-13	-11	-09	-06	-04	-02	0	02	04	07	09	12	15	17	20	24	27	30	34	38	44	49	58	65
-19	-17	-15	-13	-11	-09	-07	-04	-02	0	02	04	07	10	12	15	18	21	25	28	32	37	42	48	56	63
-21	-19	-17	-15	-13	-11	-09	-07	-04	-02	0	02	05	08	10	13	16	19	22	26	30	35	40	46	55	62
-23	-21	-19	-17	-15	-13	-11	-09	-07	-04	-02	0	02	05	08	11	13	17	20	24	28	32	38	44	53	61
-26	-24	-22	-20	-18	-16	-14	-12	-09	-07	-05	-02	0	03	06	08	11	14	18	21	26	30	36	42	51	59
-28	-26	-24	-22	-20	-18	-16	-14	-12	-10	-08	-05	-03	0	03	06	08	12	15	19	23	28	33	40	50	57
-31	-29	-27	-25	-23	-21	-19	-17	-15	-12	-10	-08	-06	-03	0	03	06	09	12	16	21	25	31	38	48	55
-33	-31	-29	-27	-26	-24	-22	-19	-17	-15	-13	-11	-08	-06	-03	0	03	06	10	13	18	23	29	36	46	53
-36	-34	-32	-30	-28	-27	-25	-23	-20	-18	-16	-13	-11	-08	-06	-03	0	03	07	11	15	20	26	33	43	51
-39	-37	-35	-33	-31	-30	-28	-26	-24	-21	-19	-17	-14	-12	-09	-06	-03	0	03	07	12	17	23	30	40	49
-42	-40	-38	-36	-34	-33	-31	-29	-27	-25	-22	-20	-18	-15	-12	-10	-07	-03	0	04	08	13	19	26	37	46
-45	-43	-41	-39	-38	-36	-34	-32	-30	-28	-26	-24	-21	-19	-16	13	-11	-07	-04	0	04	09	15	23	34	43
-48	-47	-45	-43	-41	-40	-38	-36	-34	-32	-30	-28	-26	-23	-21	-18	-15	-12	-08	-04	0	05	11	19	30	40
-52	-51	-49	-47	-45	-44	-42	-40	-38	-37	-35	-32	-30	-28	-25	-23	-20	-17	-13	-09	-05	0	06	14	25	35
-56	-55	-53	-52	-50	-48	-47	-45	-44	-42	-40	-38	-36	-33	-31	-29	-26	-23	-19	-15	-11	-06	0	08	19	30
-61	-60	-58	-57	-56	-54	-53	-51	-49	-48	-46	-44	-42	-40	-38	-36	-33	-30	-26	-23	-19	-14	-08	0	12	23
-68	-67	-66	-64	-63	-62	-61	-59	-58	-56	-55	-53	-51	-50	-48	-46	-43	-40	-37	-34	-30	-25	-19	-12	0	11
-72	-72	-71	-70	-69	-68	-67	-66	-65	-63	-62	-61	-59	-57	-55	-53	-51	-49	-46	43	-40	-35	-30	-23	-11	0

Table A-9
Height-Weight Class Division of Secondary School Girls

Height	Height		Weight		
	Feet and inches	Inches	Slender	Medium	Heavy
Short	4-7	55	Up to 82	83-111	112 up
	4-8	56	Up to 79	80-102	103 up
	4-9	57	Up to 82	83-107	108 up
	4-10	58	Up to 85	86-110	111 up
	4-11	59	Up to 91	92-112	113 up
	5-0	60	Up to 96	97-117	118 up
	5-1	61	Up to 99	100-120	121 up
Medium	5-2	62	Up to 103	104-124	125 up
	5-3	63	Up to 106	107-127	128 up
	5-4	64	Up to 110	111-132	133 up
Tall	5-5	65	Up to 113	114-136	137 up
	5-6	66	Up to 116	117-140	141 up
	5-7	67	Up to 120	121-143	144 up
	5-8	68	Up to 120	121-152	153 up
	5-9	69	Up to 126	127-153	154 up
	5-10	70	Up to 125	126-154	155 up
	5-11 up	71 up	Up to 124	125-160	161 up

Direction: Read down the height column until the girl's height is reached. Next read to the right to the weight column containing the girl's weight for that height. Her classification is the height group and weight group combined. (Example: A girl whose height is 65 inches and whose weight is 137 is a Tall Heavy [TH].)

Code for classification:

SS—Short slender	MS—Medium slender	TS—Tall slender
SM—Short medium	MM—Medium medium	TM—Tall medium
SH—Short heavy	MH—Medium heavy	TH—Tall heavy

Courtesy of C. Wesley Dane, Instructor, Indiana University. Done under the direction of Karl W. Bookwalter, Professor, Indiana University. Data (19,000 cases) furnished by Research Committee of the National Section on Women's Athletics, American Association for Health, Physical Education, and Recreation.

Table A-10
Hyde Archery Test for Women*

Scale	First Columbia total score (target score)	Scale	First Columbia total score (target score)	Scale	First Columbia total score (target score)
100	436	65	226	30	81
99	430	64	220	29	78
98	424	63	214	28	76
97	418	62	208	27	73
96	412	61	202	26	70
95	406	60	196	25	67
94	400	59	190	24	65
93	394	58	184	23	62
92	388	57	178	22	59
91	382	56	172	21	56
90	376	55	166	20	54
89	370	54	160	19	51
88	364	53	154	18	48
87	358	52	148	17	45
86	352	51	142	16	43
85	346	50	136	15	40
84	340	49	133	14	37
83	334	48	131	13	34
82	328	47	128	12	32
81	322	46	125	11	29
80	316	45	122	10	26
79	310	44	120	9	23
78	304	43	117	8	21
77	298	42	114	7	18
76	292	41	111	6	15
75	286	40	109	5	12
74	280	39	106	4	10
73	274	38	103	3	7
72	268	37	100	2	4
71	262	36	98	1	1
70	256	35	95		
69	250	34	92		
68	244	33	89		
67	238	32	87		
66	232	31	84		

* First Columbia round only. Reproduced from Edith I. Hyde, "An Achievement Scale in Archery," *The Research Quarterly*, 7:109–115, May, 1937.

Table A-11
Indiana Motor Fitness Achievement Scales for Boys

Scale score		Up to 779	780–814	815–849	850–884	885–919	920 and up	Total	Scale score
Superior	100	167	180	186	211	214	214	203	100
	95–9	159–66	171–79	177–85	200–10	203–13	203–13	192–02	95–99
	90–4	151–58	162–70	168–76	189–99	192–02	192–02	182–91	90–94
	85–9	143–50	152–61	158–67	178–88	181–91	181–91	171–81	85–89
	80–4	135–42	143–51	149–57	167–77	170–80	171–80	161–70	80–84
	75–9	127–34	134–42	140–48	156–66	158–69	160–70	151–60	75–79
Good	70–4	119–26	124–33	131–39	146–55	147–57	149–59	140–50	70–74
	65–9	112–18	113–23	122–30	135–45	136–46	138–48	130–39	65–69
	60–4	104–11	106–14	112–21	124–34	125–35	127–37	119–29	60–64
Average	55–9	96–03	96–05	103–11	113–23	114–24	117–26	109–18	55–59
	50–4	88–95	87–95	94–02	102–12	103–13	106–16	99–08	50–54
	45–9	80–87	78–86	85–93	91–01	91–02	95–05	88–98	45–49
	40–4	72–79	69–77	76–84	80–90	80–90	84–94	78–87	40–44
Poor	35–9	64–71	59–68	66–75	69–79	69–79	73–83	67–77	35–39
	30–4	56–63	50–58	57–65	58–68	58–68	63–72	57–66	30–34
	25–9	48–55	41–49	48–56	47–57	47–57	52–62	47–56	25–29
Inferior	20–4	40–47	31–40	39–47	36–46	36–46	41–51	36–46	20–24
	15–9	33–39	22–30	30–38	25–35	25–35	30–40	26–35	15–19
	10–4	25–32	13–21	20–29	14–24	13–24	19–29	15–25	10–14
	5–9	17–24	3–12	11–19	3–13	2–12	9–18	5–14	5–9
	0–4	9–16	0–2	2–10	0–2	0–2	0–8	0–4	0–4

SOURCE: State of Indiana, *Physical Fitness Manual for High School Boys,* Department of Public Instruction, Indianapolis, Ind., Bulletin no. 136, 1944, pp. 5–13.

Table A-12
Indiana Motor Fitness Achievement Scales for Girls

Scale score		SS	SM	SH	MS	MM	MH	TS	TM	TH	Total	Scale score
	100	165	177	142	149	174	149	172	150	181	174	100
	95–9	155–64	166–76	134–41	140–48	163–73	139–48	161–71	141–49	169–80	163–73	95–99
Superior	90–4	145–54	156–65	125–33	131–39	153–62	130–38	151–60	132–40	156–68	152–62	90–94
	85–9	135–44	145–55	116–24	122–30	142–52	120–29	141–50	123–31	144–55	142–51	85–89
	80–4	125–34	134–44	107–15	113–21	132–41	111–19	130–40	114–22	131–43	130–41	80–84
	75–9	115–24	123–33	99–06	104–12	121–31	102–10	120–29	105–13	118–30	120–29	75–79
	70–4	105–14	112–22	90–98	95–03	110–20	92–01	110–19	96–04	106–17	110–19	70–74
Good	65–9	95–04	101–11	81–89	86–94	100–09	83–91	99–09	88–95	93–05	99–09	65–69
	60–4	85–94	90–00	72–80	77–85	89–99	73–82	89–98	79–87	81–92	88–98	60–64
	55–9	75–84	80–89	64–71	68–76	79–88	64–72	79–88	70–78	68–80	78–87	55–59
Average	50–4	65–74	69–79	55–63	59–67	68–78	55–63	69–78	61–69	56–67	67–77	50–54
	45–9	55–64	58–68	46–54	50–58	57–67	45–54	58–66	52–60	43–55	56–66	45–49
	40–4	45–54	47–57	37–45	41–49	47–56	36–44	48–57	43–51	31–42	45–55	40–44
	35–9	35–44	36–46	29–36	32–40	36–46	26–35	38–47	34–42	18–30	35–44	35–39
Poor	30–4	25–34	25–35	20–28	23–31	26–36	17–25	27–37	27–33	6–17	24–34	30–34
	25–9	15–24	14–24	11–19	14–22	15–25	8–16	17–26	18–26	0–5	13–23	25–29
	20–4	5–14	4–13	2–10	5–13	4–14	0–7	7–16	9–17	......	3–12	20–24
	15–9	0–4	0–3	0–1	0–4	0–3	......	0–6	0–8	......	0–2	15–19
Inferior	10–4	......	......	......	......	......	......	......	......	......	......	10–14
	5–9	......	......	......	......	......	......	......	......	......	......	5–9
	0–4	......	......	......	......	......	......	......	......	......	......	0–4

SOURCE: State of Indiana, *Physical Fitness Manual for High School Girls*, Department of Public Instruction, Indianapolis, Ind., Bulletin no. 137 (revised), 1944.

Table A-13
Indiana Physical Fitness Test State Norms for Elementary Boys
(Grades Four to Eight)

	Scale score	Up to 609	610–644	645–679	680–714	715–749	750 and over	Scale score
Excellent	100	109	124	124	140	145	159	100
	99	108	122	122	138	144	158	99
	98	106	121	121	136	142	156	98
	97	105	119	119	135	140	154	97
	96	104	118	118	133	139	152	96
	95	102	116	116	132	137	151	95
	94	101	115	115	130	135	149	94
	93	100	113	113	128	134	147	93
	92	99	111	112	127	132	145	92
	91	97	110	111	125	130	144	91
	90	96	108	109	123	129	142	90
	89	95	107	108	122	127	140	89
	88	93	105	106	120	125	138	88
	87	92	104	105	118	124	136	87
	86	91	102	103	117	122	135	86
	85	90	101	102	115	120	133	85
	84	88	99	100	113	119	131	84
	83	87	98	99	112	117	129	83
	82	86	96	97	110	115	128	82
	81	85	95	96	108	114	126	81
Good	80	83	93	94	107	112	124	80
	79	82	92	93	105	110	122	79
	78	81	90	92	103	109	121	78
	77	79	89	90	102	107	119	77
	76	78	87	89	100	105	117	76
	75	77	86	87	99	104	115	75
	74	76	84	86	97	102	114	74
	73	74	83	84	95	100	112	73
	72	73	81	83	94	99	110	72
	71	72	80	81	92	97	108	71
	70	70	78	80	90	95	107	70
	69	69	77	78	89	94	105	69
	68	68	75	77	87	92	103	68
	67	67	74	76	85	90	101	67
	66	65	72	74	84	89	100	66
	65	64	71	73	82	87	98	65
	64	63	69	71	80	85	96	64
	63	61	68	70	79	84	94	63
	62	60	66	68	77	82	92	62
	61	59	65	67	75	80	91	61
Fair	60	58	63	65	74	79	89	60
	59	56	62	64	72	77	87	59
	58	55	60	62	70	75	85	58

Table A-13 (*Continued*)

Scale score		Up to 609	610–644	645–679	680–714	715–749	750 and over	Scale score
	57	54	59	61	69	74	84	57
	56	53	57	59	67	72	82	56
	55	51	56	58	66	70	80	55
	54	50	54	57	64	69	78	54
	53	49	53	55	62	67	77	53
	52	47	51	54	61	65	75	52
	51	46	50	52	59	64	73	51
Fair	50	45	48	51	57	62	71	50
	49	44	47	49	56	60	70	49
	48	42	45	48	54	59	68	48
	47	41	44	46	52	57	66	47
	46	40	42	45	51	55	64	46
	45	38	41	43	49	54	63	45
	44	37	39	42	47	52	61	44
	43	36	38	40	46	50	59	43
	42	35	36	39	44	49	57	42
	41	33	34	38	42	45	54	41
	40	32	33	36	41	44	52	40
	39	31	31	35	39	42	50	39
	38	29	20	33	37	40	48	38
	37	28	28	32	36	39	47	37
	36	27	27	30	34	37	45	36
	35	26	25	29	33	35	43	35
	34	24	24	27	31	34	41	34
	33	23	22	26	29	32	40	33
	32	22	21	24	28	30	38	32
Poor	31	21	19	23	26	29	36	31
	30	19	18	21	24	27	34	30
	29	18	16	20	23	25	33	29
	28	17	15	19	21	24	31	28
	27	15	13	17	19	22	29	27
	26	14	12	16	18	20	27	26
	25	13	10	14	16	19	26	25
	24	12	9	13	14	17	24	24
	23	10	7	11	13	15	22	23
	22	9	6	10	11	14	20	22
	21	8	4	8	9	12	19	21
	20	6	3	7	8	10	17	20
	19	5	1	5	6	9	15	19
	18	4	...	4	4	7	13	18
Inferior	17	3	...	3	3	5	12	17
	16	1	...	1	1	3	19	16
	15	...	...	...	...	2	8	15
	14	...	...	...	..	1	6	14
	13	...	...	...	...	...	4	13

Table A-13 (*Continued*)

Scale score		Up to 609	610–644	645–679	680–714	715–749	750 and over	Scale score
	12	...	...	...	...	...	3	12
	11	...	...	...	...	...	1	11
	10	...	...	...	...	...	...	10
	9	...	...	...	...	...	...	9
	8	...	...	...	...	...	...	8
Inferior	7	...	...	...	...	...	...	7
	6	...	...	...	...	...	...	6
	5	...	...	...	...	...	...	5
	4	...	...	...	...	...	...	4
	3	...	...	...	...	...	...	3
	2	...	...	...	...	...	...	2
	1	...	...	...	...	...	...	1

SOURCE: C. C. Franklin and N. G. Lehsten, "Indiana Physical Fitness Tests for the Elementary Level (Grades 4–8)," *The Physical Educator,* 5:38–45, May, 1948.

Table A-14
Indiana Physical Fitness Test State Norms for Elementary Girls
(Grades Four to Eight)

	Scale score	Up to 609	610–644	645–679	680–714	715–749	750 and over	Scale score
	100*	110	121	146	154	129	105	100
	99	108	119	144	152	128	104	99
	98	107	118	142	150	126	103	98
	97	106	116	141	148	124	102	97
	96	104	115	139	146	123	100	96
	95	103	114	137	144	121	99	95
	94	101	112	136	142	120	98	94
	93	100	111	134	140	118	97	93
	92	98	109	132	138	116	96	92
Excellent	91	97	108	131	136	115	94	91
	90	96	106	129	134	113	93	90
	89	94	105	127	132	112	92	89
	88	93	103	126	130	110	91	88
	87	91	102	124	128	109	89	87
	86	90	101	122	126	107	88	86
	85	88	99	121	124	105	87	85
	84	87	98	119	122	104	85	84
	83	86	96	117	120	102	84	83
	82	84	95	116	118	101	83	82
	81	83	93	114	116	99	82	81
	80	81	92	112	114	98	81	80
	79	80	91	111	112	96	80	79
	78	79	89	109	110	94	78	78
	77	77	88	107	108	93	77	77
	76	76	86	106	105	91	76	76
	75	74	85	104	103	90	75	75
	74	73	83	102	101	88	73	74
	73	71	82	101	99	86	72	73
	72	70	81	99	97	85	71	72
Good	71	69	79	97	95	83	70	71
	70	67	78	96	93	82	68	70
	69	66	76	94	91	80	67	69
	68	64	75	92	89	79	66	68
	67	63	73	91	87	77	65	67
	66	61	72	89	85	75	64	66
	65	60	71	87	83	74	62	65
	64	59	69	86	81	72	61	64
	63	57	68	84	79	71	60	63
	62	56	66	82	77	69	59	62
	61	54	65	81	75	68	57	61
	60	53	63	79	73	66	56	60
Fair	59	52	62	77	71	64	55	59
	58	50	61	76	69	63	54	58

Table A-14 (*Continued*)

Scale score		Up to 609	610–644	645–679	680–714	715–749	750 and over	Scale score
	57	49	59	74	67	61	53	57
	56	47	58	72	65	60	51	56
	55	46	56	71	63	58	50	55
	54	44	55	69	61	56	49	54
	53	43	53	67	59	55	48	53
	52	42	52	66	57	53	46	52
	51	40	51	64	55	52	45	51
Fair	50	39	49	63	53	50	44	50
	49	37	48	61	51	49	43	49
	48	36	46	59	49	47	41	48
	47	35	45	58	47	45	40	47
	46	33	43	56	45	44	39	46
	45	32	42	54	43	42	38	45
	44	30	41	53	41	41	37	44
	43	29	39	51	38	39	35	43
	42	27	38	49	36	37	34	42
	41	26	36	48	34	36	33	41
	40	25	35	46	32	34	32	40
	39	23	33	44	30	33	30	39
	38	22	32	43	28	31	29	38
	37	20	31	41	26	30	28	37
	36	19	29	39	24	28	27	36
	35	17	28	38	22	26	25	35
	34	16	26	36	20	25	24	34
	33	15	25	34	18	23	23	33
	32	13	23	33	16	22	22	32
Poor	31	12	22	31	14	20	21	31
	30	10	21	29	12	19	19	30
	29	9	19	28	10	17	18	29
	28	8	18	26	8	15	17	28
	27	6	16	24	6	14	16	27
	26	5	15	23	4	12	14	26
	25	3	13	21	2	11	13	25
	24	2	12	19	1	9	12	24
	23	...	11	18	...	7	11	23
	22	...	9	16	...	6	9	22
	21	...	8	14	...	4	8	21
	20	...	6	13	...	3	7	20
	19	...	5	11	...	1	6	19
	18	...	3	9	...	...	5	18
	17	...	2	8	...	...	3	17
Inferior	16	...	1	6	...	...	2	16
	15	...	...	4	...	...	1	15
	14	...	...	3	...	...	...	14
	13	...	...	1	...	...	...	13

Table A-14 (*Continued*)

Scale score		Up to 609	610–644	645–679	680–714	715–749	750 and over	Scale score
	12	...	...	...	...	...	...	12
	11	.	...	...	...	...	...	11
	10	...	...	...	...	...	...	10
	9	...	...	...	...	...	...	9
	8	...	...	...	...	...	...	8
Inferior	7	...	...	. .	...	...	...	7
	6	. .	. .	...	...	...	...	6
	5	...	...	...	...	...	...	5
	4	...	...	. .	...	...	...	4
	3	...	...	...	...	. .	...	3
	2	...	...	. .	. .	...	...	2
	1	...	...	...	. .	...	...	1

SOURCE: C. C. Franklin and N. G. Lehsten, "Indiana Physical Fitness Tests for the Elementary Level (Grades 4–8)," *The Physical Educator,* 5:38–45, May, 1948.

Table A-15

Larson's Motor Ability Tests—Outdoors

Baseball throw (wt. = 2.23)								
Raw score (feet)	T score	Wtd T score	Raw score (feet)	T score	Wtd T score	Raw score (feet)	T score	Wtd T score
270	80	178	196	56.8	127	122	32.6	73
268	79.6	178	194	56.2	125	120	32	71
266	79.2	177	192	55.6	124	118	31.6	70
264	78.8	176	190	55	123	116	31.2	70
262	78.4	175	188	54.2	121	114	30.8	69
260	78	174	186	53.4	119	112	30.4	68
258	77.8	173	184	52.6	118	110	30	67
256	77.6	173	182	51.8	116	108	29.6	66
254	77.4	173	180	51	114	106	29.2	65
252	77.2	172	178	50.2	112	104	28.8	64
250	77	172	176	49.4	110	102	28.4	63
248	76	169	174	48.6	108	100	28	62
246	75	167	172	47.8	107	98	27.4	61
244	74	165	170	47	105	96	26.8	60
242	73	163	168	46.4	103	94	26.2	58
240	72	161	166	45.8	102	92	25.6	57
238	71.2	159	164	45.2	101	90	25	56
236	70.4	157	162	44.6	99	88	24.8	55
234	69.6	155	160	44	98	86	24.6	55
232	68.8	153	158	43.4	97	84	24.4	54
230	68	152	156	42.8	95	82	24.2	54
228	67.2	150	154	42.2	94	80	24	54
226	66.4	148	152	41.6	93	78	23	51
224	65.6	146	150	41	91	76	22	49
222	64.8	145	148	40.4	90	74	21	47
220	64	143	146	39.8	89	72	20	45
218	63.2	141	144	39.2	87	70	19	42
216	62.4	139	142	38.6	86			
214	61.6	137	140	38	85			
212	60.8	136	138	37.4	83			
210	60	134	136	36.8	82			
208	59.6	133	134	36.2	81			
206	49.2	132	132	35.6	79			
204	58.8	131	130	35	78			
202	58.4	130	128	34.4	77			
200	58.0	129	126	33.8	75			
198	57.4	128	124	33.2	74			

Chinning (wt. = 3.66)		
Raw score	T score	Wtd T score
23	80	293
22	78	285
21	77	282
20	76	278
19	75	275
18	73	267
17	71	260
16	69	253
15	67	245
14	65	238
13	63	231
12	60	220
11	58	212
10	55	201
9	52	190
8	50	183
7	48	176
6	45	165
5	42	154
4	40	146
3	37	135
2	34	124
1	32	117
0	27	99

Bar-snap (wt. = 4.91)					
Raw score (inches)	T score	Wtd T score	Raw score	T score	Wtd T score
102	80	393	65	59.3	291
101	79.3	389	64	58.7	288
100	78.7	386	63	58	285
99	78	383	62	57.3	281
98	77.3	380	61	56.7	278
97	76.7	377	60	56	275
96	76	373	59	55.3	272
95	75.5	371	58	54.7	269
94	75	368	57	54	265
93	74.5	366	56	53.3	262
92	74	363	55	52.7	259
91	73.5	361	54	52	255
90	73	358	53	51.2	251
89	72.5	356	52	50.3	247
88	72	354	51	49.5	243
87	71.5	351	50	48.7	239
86	71	349	49	47.8	235
85	70.5	346	48	47	231
84	70	344	47	46	226
83	69.5	341	46	45	221
82	69	339	45	44	216
81	68.5	336	44	43	211
80	68	334	43	42	206
79	67.5	331	42	41	201
78	67	329	41	39.7	195
77	66.5	327	40	38.3	188
76	66	324	39	37	182
75	65.5	322	38	35.7	175
74	65	319	37	34.3	168
73	64.5	317	36	33	162
72	64	314	35	31.5	156
71	63.3	311	34	30	147
70	62.7	308	33	28.5	140
69	62	304	32	27	133
68	61.3	301	31	25.5	125
67	60.7	298	30	24	118
66	60	295			

Vertical jump (wt. = 1.00)		
Raw score (inches)	T score	Wtd T score
26.0	78	78
25.5	76	76
25.0	74	74
24.5	72	72
24.0	70	70
23.5	68	68
23.0	66	66
22.5	64	64
22.0	62	62
21.5	60.5	61
21.0	59	59
20.5	57.5	58
20.0	56	56
19.5	54	54
19.0	52	52
18.5	50.5	51
18.0	49	49
17.5	47	47
17.0	45	45
16.5	43.5	44
16.0	42	42
15.5	40	40
15.0	38	38
14.5	36.5	37
14.0	35	35
13.5	33	33
13.0	31	31
12.5	29	29
12.0	27	27
11.5	25.5	26
11.0	24	24

SOURCE: Leonard A. Larson, "A Factor Analysis of Motor Ability Variables and Tests, with Tests for College Men," *The Research Quarterly*, 12:499–517, October, 1941.

Table A-16
Larson's Strength Test Scoring Table

Vertical jump		Dips*		Chinning	
Raw score	Std. × 2.27	Raw score	Std. × 1.00	Raw score	Std. × 1.00
26.0	227	30	109	30	276
25.8	225	29	106	29	267
25.6	222	28	104	28	259
25.4	220	27	101	27	252
25.2	216	26	98	26	244
25.0	213	25	95	25	235
24.8	211	24	92	24	227
24.6	209	23	89	23	220
24.4	207	22	87	22	212
24.2	204	21	84	21	204
24.0	202	20	81	20	197
23.8	200	19	78	19	189
23.6	197	18	75	18	180
23.4	195	17	72	17	174
23.2	193	16	70	16	165
23.0	188	15	67	15	157
22.8	186	14	64	14	148
22.6	184	13	61	13	142
22.4	182	12	58	12	134
22.2	179	11	55	11	125
22.0	177	10	52	10	119
21.8	175	9	50	9	110
21.6	173	8	47	8	102
21.4	170	7	44	7	95
21.2	168	6	41	6	87
21.0	166	5	38	5	78
20.8	161	4	35	4	72
20.6	159	3	33	3	64
20.4	157	2	30	2	55
20.2	154	1	27	1	47
20.0	152				
19.8	150				
19.6	148				
19.4	145				
19.2	143				
19.0	141				
18.8	138				
18.6	134				
18.4	132				
18.2	129				
18.0	127				
17.8	125				
17.6	123				

Table A-16 (*Continued*)

Vertical jump		Dips		Chinning	
Raw score	Std. × 2.27	Raw score	Std. × 1.00	Raw score	Std. × 1.00
17.4	120				
17.2	118				
17.0	116				
16.8	114				
16.6	111				
16.4	107				
16.2	104				
16.0	102				
15.8	100				
15.6	98				
15.4	95				
15.2	93				
15.0	91				
14.8	89				
14.6	86				
14.4	84				
14.2	79				
14.0	77				
13.8	75				
13.6	73				
13.4	70				
13.2	68				
13.0	66				
12.8	64				
12.6	61				
12 4	59				
12.2	57				
12.0	54				
11.8	50				
11.6	48				
11.4	45				
11.2	43				
11.0	41				
10.8	39				
10.6	36				
10.4	34				
10.2	32				
10.0	27				
9.8	25				
9.6	23				
9.4	20				
9.2	18				
9.0	16				
8.8	14				

Table A-16 (*Continued*)

Vertical jump		Dips		Chinning	
Raw score	Std. × 2.27	Raw score	Std. × 1.00	Raw score	Std. × 1.00
8.6	11				
8.4	9				
8.2	7				
8.0	5				

SOURCE: L. A. Larson, "A Factor and Validity Analysis of Strength Variables and Tests with a Test Combining Chinning, Dipping, and Vertical Jumps," *The Research Quarterly*, 11:82-96, December, 1940.

Table A-17
Navy Standard Physical Fitness Test T Scores

T score	Squat-thrusts	Sit-ups	Push-ups	Squat-jump	Pull-ups
100	48	205	89	127	37
99		200	86-8	123	
98	47	190-5	84-5	120	36
97		185	82-3	117	35
96		180	80-1	114	
95	46	170-5	78-9	111	34
94		165	76-7	108	33
93		160	74-5	106	32
92	45	155	73	103	31
91		150	71-2	100	30
90	44	145	69-70	98	29
89		140	67-8	95	
88		135	66	93	28
87	43	130	64-5	90	27
86		125	63	88	
85		120	61-2	86	26
84	42	115	60	83	
83			58-9	81	25
82	41	110	57	79	24
81		105	55-6	77	
80		100	54	75	23
79	40	97-9	53	73	22
78		92-6	52	71	
77		89 91	51	69	21
76	39	86-8	50	68	
75		82-5	49	66	20
74	38	79-81	48	64	19
73		77-8	47	63	
72		74-6	45-6	61	18
71	37	72-3	44	59	
70		69-71	43	58	17
69		67-8	42	56	
68	36	65-6	41	55	16
67		63-4		54	
66	35	61-2	40	52	15
65		59-60	39	51	
64		57-8	38	49	14
63	34	55-6	37	48	
62		54	36	47	13
61	33	52-3		46	
60		51	35	45	
59		49-50	34	43	12
58	32	47-8	33	42	
57		46	32	41	
56	31	44-5		40	11
55		43	31	39	
54		42	30	38	
53	30	40-1	29	37	10

T score	Squat-thrusts	Sit-ups	Push-ups	Squat-jump	Pull-ups
52		39		36	
51	29	38	28	35	
50		36-7		34	9
49	28	35	27	33	
48		34	26		
47		33		32	8
46	27	32	25	31	
45		31		30	
44	26	29-30	24	29	
43		28	23		7
42	25	27		28	
41		26		27	
40		25	22		6
39	24		21	26	
38		24		25	
37	23				5
36		23	20	24	
35	22	22		23	
34			19		4
33	21	21	18	22	
32		20			
31	20	19	17	21	
30					
29	19	18		20	
28		17	16	19	3
27	18	16			
26			15		
25	17			18	
24		15	14		
23	16			17	2
22		14			
21	15		13	16	
20		13			
19	14		12	15	1
18		12			
17	13				
16			11	14	
15	12	11			
14	11				
13			10	13	
12	10	10			
11			9		
10	9			12	
9	8	9			
8			8		
7	7			11	
6	6				
5	5	8	7		
4	4				
3			6	10	
2	3				
1	2	7			

SOURCE: Bureau of Naval Personnel, *Physical Fitness Manual for the U.S. Navy*, Washington, 1943, Training Division, Physical Section, chap. IV.

Table A-18
NSWA Physical Performance Test Scoring Table

Scale score	Standing broad jump	Basket-ball throw	Potato race	Pull-ups	Push-ups	Sit-ups	10-sec squat-thrust	30-sec squat-thrust	Scale score
100	7–9	78	8.4	47	61	65	9–1	24	100
95	7–7	75	8.6	45	58	61	9	23	95
90	7–4	72	8.8	42	54	57	8–3	22	90
85	7–2	68	9.0	39	51	54	8–1	21	85
80	6–11	65	9.4	37	47	50	8	20	80
75	6–9	62	9.6	34	43	46	7–3	19	75
70	6–7	59	10.0	32	39	43	7–1	18–2	70
65	6–4	56	10.2	29	36	39	7	18	65
60	6–2	53	10.4	26	32	36	6–2	17	60
55	6–0	50	10.6	24	28	33	6–1	16	55
50	5–9	46	11.0	21	25	29	6	15	50
45	5–7	43	11.2	18	21	25	5–2	14–2	45
40	5–5	40	11.6	16	17	22	5–1	14	40
35	5–2	37	11.8	13	13	18	4–3	13	35
30	5–0	34	12.0	10	10	15	4–2	12	30
25	4–9	31	12.4	8	6	11	4	11	25
20	4–7	27	12.6	5	2	7	3–3	10	20
15	4–4	24	13.0	3	1	3	3–2	9	15
10	4–2	21	13.2	1	0	1	3	8–2	10
5	4–0	18	13.4	0	0	0	2–3	7–2	5
0	3–9	15	13.6	0	0	0	2–2	7	0

SOURCE: Eleanor Metheny, "Physical Performance Levels for High School Girls," *Journal of Health and Physical Education,* 16:84, June, 1945.

Table A-19
Oregon Motor Fitness Test Norms (Boys)
Grades Four, Five, and Six

Classification	Superior	Good	Fair	Poor	Inferior	Grade
Standing broad jump	67–up	62–66	52–61	43–51	12–42	
Push-ups	36–up	25–35	10–24	3–9	0–2	4
Sit-ups	90–up	65–89	30–64	5–25	0–4	
Total points	205–up	190–204	140–189	120–139	0–119	
Standing broad jump	71–up	66–70	56–65	47–55	16–46	
Push-ups	34–up	24–33	10–23	4–6	0–3	5
Sit-ups	95–up	70–94	33–69	5–32	0–4	
Total points	207–up	192–206	142–191	122–141	0–121	
Standing broad jump	75–up	70–74	59–69	50–59	18–49	
Push-ups	33–up	24–32	12–23	6–11	0–5	6
Sit-ups	105–up	75–104	35–74	5–34	0–4	
Total points	205–up	195–204	145–194	125–144	0–124	
Grades Seven, Eight, and Nine						
Jump and reach	19–up	17–18	14–16	11–13	0–10	
Pull-up	8–up	5–7	2–4	1	0	7
Potato race	0–32	33–35	36–39	40–42	43–up	
Total points	203–up	176–202	145–175	128–144	0–127	
Jump and reach	21–up	18–20	14–17	11–13	0–10	
Pull-up	9–up	6–8	2–5	1	0	8
Potato race	0–31	32–34	35–38	39–41	42–up	
Total points	188–up	172–187	130–171	104–129	0–103	
Jump and reach	23–up	20–22	16–19	13–15	0–12	
Pull-up	11–up	8–10	4–7	1–3	0	9
Potato race	0–29	30–32	33–36	37–39	40–up	
Total points	192–up	172–191	134–171	111–133	0–110	

Table A-19 (*Continued*)
Grades Ten, Eleven, and Twelve

Classification	Superior	Good	Fair	Poor	Inferior	Grade
Jump and reach	24–up	21–23	17–20	14–16	0–13	
Pull-up	12–up	8–11	4–7	1–3	0	10
Potato race	0–30	31–32	33–35	36–38	39–up	
Total points	203–up	174–202	138–173	106–137	0–105	
Jump and reach	26–up	23–25	19–22	15–18	0–14	
Pull-up	14–up	10–13	6–9	2–5	0–1	11
Potato race	0–29	30–31	32–35	36–37	38–up	
Total points	208–up	175–207	137–174	100–136	0–99	
Jump and reach	27–up	24–26	20–23	16–19	0–15	
Pull-up	15–up	11–14	6–10	2–5	0–1	12
Potato race	0–29	30–31	32–34	35–36	37–up	
Total points	209–up	182–208	143–182	108–142	0–107	

SOURCE: Oregon State Department of Education, *Oregon Motor Fitness Test Battery,* Salem, Ore., 1959.

Table A-20
Oregon Motor Fitness Test Norms (Girls)
Grades Four, Five, and Six

Classification	Superior	Good	Fair	Poor	Inferior	Grade
Hanging in arm flexed position	45–up	30–44	15–29	1–14	0	
Standing broad jump	65–up	55–64	49–54	38–48	0–37	4
Crossed-arm curl-up	90–up	66–89	34–65	10–33	0–9	
Hanging in arm flexed position	44–up	31–43	19–30	6–18	0–5	
Standing broad jump	70–up	64–69	55–63	44–54	0–43	5
Crossed-arm curl-up	92–up	68–91	36–67	12–35	0–11	
Hanging in arm flexed position	52–up	37–51	22–36	7–21	0–6	
Standing broad jump	73–up	62–72	55–61	44–54	0–43	6
Crossed-arm curl-up	95–up	71–94	39–70	15–38	0–14	
Total points	198–up	168–197	138–167	115–137	114–down	

Grades Seven, Eight, and Nine

Classification	Superior	Good	Fair	Poor	Inferior	Grade
Hanging in arm flexed position	42–up	20–41	9–19	3–8	0–2	
Standing broad jump	75–up	66–74	57–65	48–56	0–47	7
Crossed-arm curl-up	104–up	42–103	24–41	14–23	0–13	
Hanging in arm flexed position	42–up	19–41	8–18	3–7	0–2	
Standing broad jump	77–up	68–76	59–67	50–58	0–49	8
Crossed-arm curl-up	74–up	38–73	24–37	14–23	0–13	
Hanging in arm flexed position	61–up	28–60	12–27	5–11	0–4	
Standing broad jump	80–up	71–79	61–70	52–60	0–51	9
Crossed-arm curl-up	102–up	42–101	26–41	16–25	0–15	
Total points	197–up	167–196	137–166	106–136	105–down	

Table A-20 (*Continued*)
Grades Ten, Eleven, and Twelve

Classification	Superior	Good	Fair	Poor	Inferior	Grade
Hanging in arm flexed position	50-up	24-49	10-23	3-9	0-2	
Standing broad jump	80-up	70-79	61-69	52-60	0-51	10
Crossed-arm curl-up	80-up	38-79	26-37	16-25	0-15	
Hanging in arm flexed position	58-up	25-57	10-24	4-9	0-3	
Standing broad jump	81-up	71-80	62-70	53-61	0-52	11
Crossed-arm curl-up	80-up	36-79	24-35	14-23	0-13	
Hanging in arm flexed position	46-up	26-45	11-25	3-10	0-2	
Standing broad jump	79-up	70-78	63-69	53-62	0-52	12
Crossed-arm curl-up	74-up	42-73	26-41	16-25	0-15	
Total points	197-up	167-196	137-166	106-136	105-down	

SOURCE: Oregon State Department of Education, *Oregon Motor Fitness Test Battery,* Salem, Ore., 1959.

Table A-21
Percentile Scale—Cowell Social Adjustment Index

Raw score	Percentile score	Raw score	Percentile score	Raw score	Percentile score
88	99.55	36	68.47	−6	25.22
81	99.10	35	67.12	−7	23.87
80	98.65	34	65.32	−8	23.42
79	98.20	33	64.41	−9	22.97
78	97.75	32	63.51	−12	22.52
77	97.30	31	61.26	−15	21.62
75	96.85	30	59.91	−16	21.17
74	96.40	29	59.01	−17	20.72
73	95.94	28	57.21	−18	19.82
72	95.50	27	56.31	−19	18.92
70	95.04	26	55.40	−20	18.47
68	94.59	25	54.95	−21	17.51
65	92.79	24	53.60	−23	16.22
63	92.34	23	52.70	−25	15.32
62	91.44	22	51.80	−26	13.96
61	90.54	21	50.90	−27	12.36
60	90.09	20	50.45	−28	12.16
59	89.19	18	48.65	−29	11.71
58	88.29	17	46.85	−35	10.81
57	86.49	16	45.94	−36	9.91
56	86.04	15	45.50	−39	8.56
55	85.14	14	45.04	−40	8.11
54	84.23	13	43.24	−42	7.66
52	83.33	12	41.44	−43	7.21
51	82.88	11	40.54	−44	6.76
50	82.43	10	40.09	−45	6.31
49	81.08	9	38.74	−46	5.40
48	80.18	8	37.84	−47	4.50
47	79.28	7	36.94	− 9	4.05
46	78.38	6	35.59	−50	3.60
45	77.03	4	34.68	−54	3.15
44	74.77	3	33.33	−55	2.70
43	73.42	2	32.43	−58	1.80
42	72.97	1	30 63	−61	1.35
41	72.52	−1	29.73	−62	.90
40	72.07	−2	28.38	−71	.45
38	71.17	−3	27.48	−73	.00
37	69.37	−5	26.00		

SOURCE: Charles C. Cowell, "Validating an Index of Social Adjustment for High School Use," *The Research Quarterly,* 29:7–18, March, 1958.

Table A-22
Russell-Lange Volleyball Test for Junior High School Girls (Sigma-scale Values)

Sigma scale	Test scores: Serve	Test scores: Repeated volleys	Sigma scale	Test scores: Serve	Test scores: Repeated volleys
100	—	51	50	—	22
99	45	50	49	16	—
98	44	—	48	—	21
97	—	49	47	—	—
96	43	—	46	15	20
95	—	48	45	—	—
94	42	—	44	—	19
93	41	47	43	14	—
92	—	46	42	—	—
91	40	—	41	—	18
90	—	45	40	13	—
89	39	—	39		17
88	—	44	38		—
87	38	43	37	12	16
86	37	—	36	—	—
85	—	42	35	—	15
84	36	—	34	11	—
83	—	41	33	—	—
82	35	—	32		14
81	—	40	31	10	—
80	34	39	30	—	13
79	33	—	29	—	—
78		38	28	9	12
77	32	—	27	—	—
76	—	37	26	—	—
75	31	36	25	8	11
74	30	—	24	—	—
73	—	35	23	—	10
72	29	—	22	7	—
71	—	34	21	—	9
70	28	—	20	—	—
69	—	33	19	6	8
68	27	32	18	—	—
67	26	—	17	—	—
66	—	31	16	—	7
65	25	—	15	—	—
64	—	30	14	—	6
63	24	—	13	4	—
62	23	29	12	—	5
61	—	28	11	—	—
60	22	—	10	3	4
59	—	27	9	—	—
58	21	—	8	—	—
57	—	26	7	2	3
56	20	25	6	—	—
55	19	—	5	—	2
54	—	24	4	1	—
53	18	—	3	—	1
52	—	23	2	—	—
51	17	—	1	—	—

SOURCE: Naomi Russell and Elizabeth Lange, "Achievement Tests in Volleyball for Junior High School Girls," *The Research Quarterly*, 11:33–35, December, 1940.

Table A-23
Schneider Cardiovascular Test Scoring Table

A. Reclining pulse rate		*B.* Pulse rate increase on standing				
Rate	Points	0–10 beats, points	11–18 beats, points	19–26 beats, points	27–34 beats, points	35–42 beats, points
50–60	3	3	3	2	1	0
61–70	3	3	2	1	0	−1
71–80	2	3	2	0	−1	−2
81–90	1	2	1	−1	−2	−3
91–100	0	1	0	−2	−3	−3
101–110	−1	0	−1	−3	−3	−3

C. Standing pulse rate		*D.* Pulse rate increase immediately after exercise				
Rate	Points	0–10 beats, points	11–20 beats, points	21–30 beats, points	31–40 beats, points	41–50 beats, points
60–70	3	3	3	2	1	0
71–80	3	3	2	1	0	0
81–90	2	3	2	1	0	−1
91–100	1	2	1	0	−1	−2
101–110	1	1	0	−1	−2	−3
111–120	0	1	−1	−2	−3	−3
121–130	0	0	−2	−3	−3	−3
131–140	−1	0	−3	−3	−3	−3

E. Return of pulse rate to standing normal after exercise		*F.* Systolic pressure, standing, compared with reclining	
Seconds	Points	Change (in millimeters)	Points
0–30	3	Rise of 8 or more	3
31–60	2	Rise of 2–7	2
61–90	1	No rise	1
91–120	0	Fall of 2–5	0
After 120: 2–10 beats above normal	−1	Fall of 6 or more	−1
After 120: 11–30 beats above normal	−2		

SOURCE: E. C. Schneider, "A Cardiovascular Rating as a Measure of Physical Fatigue and Efficiency" *Journal of the American Medical Association,* 74:1507–1509, May 29, 1920.

Table A-24
Scott Motor Ability Test for High School Girls (T Scales)

T score	Wall pass (410) *	Basketball throw (feet) (310) *	Broad jump (inches) (287) *	4-sec dash (yards) (398) *	Obstacle race (seconds) (374) *
80	16	71			
79			96		
78					
77	15	68	94	27	
76		66			18.5-18.9
75		65			
74		64	92		
73	14	63			
72		61			
71		59	90	26	
70		55	88		19.0-19.4
69	13	54			
68		52	86	25	
67		51			19.5-19.9
66		50			
65		49			
64		48	84	24	20.0 20 4
63	12	47			
62		46	82		20.5-20.9
61			80		
60		45	23		
59		44	78		21.0-21.4
58	11	43			
57		42	76		21.5 21.9
56		41			
55		40	74	22	
54					22.0-22.4
53		39			
52	10		72		
51		37			22 5 22.9
50		36		21	
49		35	70		
48			68		23.0 23.4
47		34	66		
46	9	33			23.5-23.9
45		32	64	20	
44		31			24.0 24.4
43			62		
42		30			24.5-24.9
41	8	29	60	19	
40		28			
39			58		25.0-25.4
38		27	56		
37	7		54		25.5-25.9

Table A-24 (*Continued*)

T score	Wall pass (410)*	Basketball throw (feet) (310)*	Broad jump (inches) (287)*	4-sec dash (yards) (398)*	Obstacle race (seconds) (374)*
36		26			26.0-26.4
35			52	18	26.5-26.9
34		25	50		27.0-27.4
33					
32		24	47		27.5-27.9
31	6	23			
30			44		
29		22		17	28.5-28.9
28					29.0-29.4
27		21			29.5-29.9
26			40		30.0-30.4
25	5	20			
24				16	30.5-30.4
23		19	36		31.5-32.4
22				15	32.5-34.9
21		16			
20	4			14	35.0-36.0

* Figure in parentheses indicates the number of subjects on which the scale is based.

SOURCE: M. Gladys Scott and Esther French, *Evaluation in Physical Education*, St. Louis, The C. V. Mosby Company, 1950, pp. 200-202.

Table A-25
Stroup Basketball Test Scale Scores

Shooting	Passing	Dribbling	Scale score	Shooting	Passing	Dribbling	Scale score
6	53	27	51	24	78	42	76
7	55		52				77
8	56	28	53	25	79	43	78
9	57	29	54	26	80		79
	59	30	55	27	81	44	80
10	60	31	56		82		81
11	61		57	28		45	82
12	62	32	58	29	83		83
13	64	33	59		84	46	84
14	65	34	60	30	85		85
	66		61		86	47	86
15		35	62	31	87		87
16	67		63	32	88	48	88
	68	36	64		89	49	89
17	69		65	33	90	50	90
	70	37	66	34	91		91
18			67	35	93	51	92
19	71	38	68	36	94		83
	72		69	37	95	52	94
20	73	39	70		97		95
21			71	38	98	53	96
	74	40	72	39	99		97
22	75		73	40	100	54	98
23	76	41	74	41	102	55	99
	77		75	42	103	56	100

SOURCE: Frances Stroup, "Game Results as a Criterion for Validating Basketball Skill Test," *The Research Quarterly,* 26:353-357, October, 1955.

Table A-26
Strength Index Norms, Boys (Belt)

AGE / AGE / WEIGHT

	8	8-6	9	9-6	10	10-6	11	11-6	12	12-6	13	13-6	14	14-6	15	15-6	16	16-6	17	17-6	18	
180															2664	2813	2917	2993	3056	3105	3159	180
178															2682	2778	2880	2954	3016	3064	3118	178
176															2601	2743	2843	2916	2976	3023	3077	176
174														2363	2569	2708	2805	2877	2936	2983	3035	174
172														2336	2537	2674	2768	2838	2896	2942	2994	172
170														2308	2505	2639	2731	2799	2856	2901	2958	170
168														2280	2474	2604	2694	2761	2816	2860	2911	168
166													2067	2253	2442	2569	2657	2722	2776	2819	2870	166
164													2043	2225	2410	2534	2620	2683	2736	2779	2829	164
162													2020	2198	2378	2499	2583	2645	2696	2738	2787	162
160												1848	1996	2170	2346	2465	2546	2606	2656	2697	2746	160
158												1827	1972	2142	2315	2430	2509	2567	2616	2656	2704	158
156											1718	1807	1949	2115	2283	2395	2472	2528	2576	2615	2663	156
154											1699	1786	1925	2087	2254	2360	2435	2490	2536	2575	2622	154
152											1680	1766	1902	2060	2219	2325	2398	2451	2496	2534	2580	152
150										1610	1661	1745	1878	2032	2188	2291	2361	2412	2456	2493	2539	150
148										1591	1642	1724	1854	2004	2156	2256	2324	2373	2416	2452	2498	148
146										1573	1623	1704	1831	1977	2124	2221	2287	2335	2376	2411	2456	146
144									1525	1554	1604	1683	1807	1949	2092	2186	2250	2296	2336	2371	2415	144
142									1506	1536	1585	1663	1784	1922	2060	2151	2213	2257	2296	2330	2374	142
140								1470	1487	1517	1566	1642	1760	1894	2029	2116	2176	2219	2256	2289	2332	140
138								1451	1468	1498	1547	1621	1736	1866	1997	2082	2139	2180	2216	2248	2291	138
136							1401	1431	1449	1480	1528	1601	1713	1839	1965	2047	2102	2141	2176	2207	2249	136
134							1382	1412	1430	1461	1509	1580	1689	1811	1933	2012	2065	2102	2136	2167	2208	134
132							1362	1392	1411	1443	1490	1560	1666	1784	1901	1977	2028	2064	2096	2126	2167	132
130						1318	1343	1372	1392	1424	1471	1539	1642	1756	1870	1942	1991	2025	2056	2085	2125	130
128					1273	1299	1324	1353	1373	1405	1452	1518	1618	1728	1838	1907	1954	1986	2016	2044	2084	128
126				1227	1255	1280	1304	1333	1354	1387	1433	1498	1595	1701	1806	1873	1917	1948	1976	2003	2043	126
124			1185	1209	1236	1261	1285	1314	1334	1368	1414	1477	1571	1673	1774	1838	1879	1909	1936	1963	2001	124
122		1146	1169	1192	1218	1242	1266	1294	1315	1350	1395	1457	1548	1646	1743	1803	1842	1870	1896	1922	1960	122
120		1130	1152	1174	1199	1223	1246	1275	1296	1331	1376	1436	1524	1618	1711	1768	1805	1831	1856	1881	1919	120
118	1092	1114	1135	1156	1181	1204	1227	1255	1277	1312	1357	1415	1500	1590	1679	1733	1768	1793	1816	1840	1877	118
116	1077	1098	1119	1139	1162	1185	1208	1235	1258	1294	1338	1395	1477	1563	1647	1699	1731	1754	1776	1799	1836	116
114	1061	1082	1102	1121	1144	1166	1188	1216	1239	1275	1319	1374	1453	1535	1615	1664	1694	1715	1736	1759	1795	114
112	1046	1066	1085	1104	1125	1146	1169	1196	1220	1257	1300	1354	1430	1508	1584	1629	1657	1677	1696	1718	1753	112
110	1030	1050	1069	1086	1107	1127	1150	1177	1201	1238	1281	1333	1406	1480	1552	1594	1620	1638	1656	1677	1712	110
108	1014	1034	1052	1068	1088	1108	1130	1157	1182	1219	1262	1312	1382	1452	1520	1559	1583	1599	1616	1636	1670	108
106	999	1018	1036	1051	1070	1089	1111	1138	1163	1201	1243	1292	1359	1425	1488	1524	1546	1560	1576	1595	1629	106
104	983	1002	1018	1033	1051	1070	1092	1118	1144	1188	1224	1271	1335	1397	1457	1490	1509	1522	1536	1555	1588	104
102	968	986	1002	1016	1033	1051	1072	1090	1125	1164	1205	1251	1312	1370	1425	1455	1472	1483	1496	1514	1546	102
100	952	970	985	998	1014	1032	1053	1079	1106	1145	1186	1230	1288	1342	1393	1420	1435	1444	1456	1473	1505	100
98	936	954	968	980	995	1013	1034	1059	1087	1126	1167	1209	1264	1314	1361	1385	1398	1405	1416	1432	1464	98
96	921	938	952	963	977	994	1014	1040	1068	1108	1148	1189	1241	1287	1329	1350	1361	1367	1376	1391	1422	96
94	905	922	935	945	958	975	995	1020	1049	1089	1129	1168	1217	1259	1298	1316	1324	1328	1336	1351		94
92	890	906	918	928	940	956	976	1001	1030	1071	1110	1148	1194	1232	1266	1281	1287	1289	1296	1310		92
90	874	890	902	910	921	937	956	981	1011	1052	1091	1127	1170	1204	1234	1246	1250	1251	1256			90
88	858	874	885	892	903	918	937	962	992	1033	1072	1106	1146	1176	1202	1211	1213	1212	1216			88
86	843	858	868	875	884	898	918	942	973	1015	1053	1086	1123	1149	1171	1176	1176	1173	1176			86
84	827	842	851	857	866	879	898	923	954	996	1034	1065	1099	1121	1139	1141	1139	1134	1136			84
82	812	826	835	840	847	860	879	903	935	978	1015	1045	1076	1094	1107	1107	1102	1096				82
80	796	810	818	822	829	841	860	883	916	959	996	1024	1052	1066	1075	1072	1065	1057				80
78	780	794	801	804	810	822	840	864	897	940	977	1003	1028	1038	1043	1037	1028	1018				78
76	765	778	786	787	792	803	821	844	878	922	958	983	1005	1011	1012	1002	991					76
74	749	762	768	769	773	784	802	825	858	903	939	962	981	983	980	967	953					74
72	734	746	751	752	755	765	782	805	839	885	920	942	958	956	948	933						72
70	718	730	735	734	736	746	763	786	820	866	901	921	934	928	916	898						70
68	702	714	718	716	718	727	744	766	801	847	882	900	910	900	885							68
66	687	698	701	699	699	708	724	746	782	829	863	880	887	873								66
64	671	682	684	681	681	689	705	727	763	810	844	859	863	845								64
62	656	666	668	664	662	669	686	707	744	792	825	839	840									62
60	640	650	651	646	644	650	666	688	725	773	806	818										60
58	624	634	634	628	625	631	647	668	706	754	787	797										58
56	609	618	618	611	607	612	628	649	687	736	768											56
54	593	602	601	593	588	593	608	629	668	717	749											54
52	578	586	584	576	570	574	589	610	649	699												52
50	562	570	568	558	551	555	570	590	630	680												50
	8	8-6	9	9-6	10	10-6	11	11-6	12	12-6	13	13-6	14	14-6	15	15-6	16	16-6	17	17-6	18	
*	7.80	8.00	8.35	8.80	9.26	9.54	9.67	9.78	9.52	9.30	9.50	10.30	11.80	13.80	15.89	17.41	18.52	19.86	20	20.40	20.68	

*Weight Deviation Multiplier

1. Norms for individuals whose weights are above limits for which norms are included are calculated by adding to the norm for any chosen weight the pound difference between that weight and the individual's weight times the Weight Deviation Multiplier.

Reproduced with permission of Frederick Rand Rogers.

Table A-27
Strength Index Norms, Girls (Belt)

AGE

WEIGHT	8	8-6	9	9-6	10	10-6	11	11-6	12	12-6	13	13-6	14	14-6	15	15-6	16	16-6	17	17-6	18	WEIGHT
180															2840	2931	2990	2984	2960	2912	2835	180
178														2721	2796	2885	2940	2934	2912	2865	2790	178
176													2596	2681	2755	2837	2890	2884	2863	2819	2745	176
174												2492	2560	2640	2707	2789	2840	2835	2815	2772	2700	174
172											2364	2457	2525	2600	2663	2741	2790	2783	2766	2726	2658	172
170										2250	2332	2422	2486	2559	2619	2693	2740	2735	2718	2679	2610	170
168										2220	2300	2387	2449	2518	2575	2645	2690	2685	2670	2632	2565	168
166									2085	2190	2268	2352	2418	2478	2531	2597	2640	2635	2621	2586	2520	166
164									2058	2160	2236	2318	2376	2437	2486	2549	2590	2586	2573	2539	2475	164
162								1923	2030	2130	2204	2283	2337	2397	2442	2501	2540	2536	2524	2493	2430	162
160							1818	1897	2002	2100	2172	2248	2300	2356	2398	2453	2490	2486	2476	2446	2385	160
158							1793	1871	1974	2070	2140	2213	2265	2315	2354	2405	2440	2436	2428	2399	2340	158
156						1736	1769	1845	1946	2040	2108	2178	2226	2275	2310	2357	2390	2386	2379	2353	2295	156
154						1711	1744	1818	1919	2010	2076	2144	2186	2234	2265	2309	2340	2337	2331	2306	2250	154
152					1681	1687	1720	1792	1891	1980	2044	2109	2151	2194	2221	2261	2290	2287	2282	2260	2205	152
150				1625	1657	1663	1693	1766	1863	1950	2012	2074	2114	2153	2177	2213	2240	2237	2234	2213	2160	150
148			1524	1602	1633	1639	1670	1740	1835	1920	1980	2039	2077	2112	2133	2165	2190	2187	2186	2166	2115	148
146			1502	1578	1609	1615	1646	1714	1807	1890	1948	2004	2040	2072	2089	2117	2140	2137	2137	2120	2070	146
144		1419	1479	1555	1584	1590	1621	1687	1780	1860	1916	1970	2002	2031	2044	2069	2090	2088	2089	2073	2025	144
142	1340	1398	1457	1531	1560	1566	1597	1661	1752	1830	1884	1915	1965	1991	2000	2021	2040	2038	2040	2027	1980	142
140	1319	1376	1435	1508	1536	1542	1572	1635	1724	1800	1852	1900	1928	1950	1956	1973	1990	1988	1992	1980	1935	140
138	1298	1354	1413	1485	1512	1518	1547	1609	1696	1770	1820	1865	1891	1909	1912	1925	1940	1938	1944	1933	1890	138
136	1278	1333	1391	1461	1488	1494	1523	1583	1668	1740	1788	1830	1854	1869	1868	1877	1890	1888	1895	1887	1845	136
134	1257	1311	1368	1438	1463	1469	1498	1556	1641	1710	1756	1796	1816	1828	1823	1829	1840	1839	1847	1840	1800	134
132	1237	1290	1346	1414	1439	1445	1474	1530	1613	1680	1724	1761	1779	1788	1779	1781	1790	1789	1798	1794	1755	132
130	1216	1268	1324	1391	1415	1421	1449	1504	1585	1650	1692	1726	1742	1747	1735	1733	1740	1739	1750	1747	1710	130
128	1195	1246	1302	1368	1391	1397	1424	1478	1557	1620	1660	1691	1705	1706	1691	1685	1690	1689	1702	1700	1666	128
126	1175	1225	1280	1344	1367	1373	1400	1452	1529	1590	1628	1656	1668	1666	1647	1637	1640	1639	1653	1654	1620	126
124	1154	1203	1257	1321	1342	1348	1375	1425	1502	1560	1596	1622	1630	1625	1602	1589	1590	1590	1605	1607	1575	124
122	1134	1182	1235	1297	1318	1324	1351	1399	1474	1530	1564	1587	1593	1585	1558	1541	1540	1540	1556	1561	1530	122
120	1113	1160	1213	1274	1294	1300	1326	1373	1446	1500	1532	1552	1556	1544	1514	1493	1490	1490	1508	1514	1485	120
118	1092	1138	1191	1251	1270	1276	1301	1347	1418	1470	1500	1517	1519	1503	1470	1445	1440	1440	1460	1467	1440	118
116	1072	1117	1169	1227	1246	1252	1277	1321	1390	1440	1468	1482	1482	1463	1426	1397	1390	1390	1411	1421	1395	116
114	1051	1095	1146	1204	1221	1227	1252	1294	1363	1410	1436	1448	1444	1422	1381	1349	1340	1341	1363	1374	1350	114
112	1031	1074	1124	1180	1197	1203	1228	1268	1335	1380	1404	1413	1407	1382	1337	1301	1290	1291	1314	1328	1305	112
110	1010	1052	1102	1157	1173	1179	1203	1242	1307	1350	1372	1378	1370	1341	1293	1253	1240	1241	1266	1281	1260	110
108	989	1030	1080	1134	1149	1155	1178	1216	1279	1320	1340	1343	1333	1300	1249	1205	1190	1191	1218	1234	1215	108
106	969	1009	1058	1110	1125	1131	1154	1190	1251	1290	1308	1308	1296	1260	1205	1157	1140	1141	1169	1188	1170	106
104	948	987	1035	1087	1100	1106	1129	1163	1224	1260	1276	1274	1258	1219	1160	1109	1090	1092	1121	1141	1125	104
102	928	966	1013	1063	1076	1082	1105	1137	1196	1230	1244	1239	1221	1179	1116	1061	1040	1042	1072	1095	1080	102
100	907	944	991	1040	1052	1058	1080	1111	1168	1200	1212	1204	1184	1138	1072	1013	990	992	1024	1048	1035	100
98	886	922	969	1017	1028	1034	1055	1085	1140	1170	1180	1169	1147	1097	1028	965	940	942	976	1001	990	98
96	866	901	947	993	1004	1010	1031	1059	1112	1140	1148	1134	1110	1057	984	917	890	892	927	955	945	96
94	845	879	924	970	979	985	1006	1032	1085	1110	1116	1100	1072	1016	939	869	840	843	879	908	900	94
92	825	858	902	946	955	961	982	1006	1057	1080	1084	1065	1035	976	895	821	790	793	830	862	855	92
90	804	836	880	923	931	937	957	980	1029	1050	1052	1030	998	935	851	773	740	743	782	815	810	90
88	783	814	858	900	907	913	932	954	1001	1020	1020	995	961	894	807	725	690	693	734	768	765	88
86	763	793	836	876	883	889	908	928	973	990	988	960	924	854	763	677	640	643	685	722	720	86
84	742	771	813	853	858	864	883	901	946	960	956	926	886	813	718	629	590	594	637	675	675	84
82	722	750	791	829	834	810	859	875	918	930	924	891	849	773	674	581	540	544	588	629	630	82
80	701	728	769	806	810	816	834	849	890	900	892	856	812	732	630	533	490	494	540	582	585	80
78	680	706	747	783	786	792	809	823	862	870	860	821	775	691	586	485	440	444	492	535		78
76	660	685	725	759	762	768	785	797	834	840	828	786	738	651	542	437	390	394	443	489		76
74	639	663	702	736	737	743	760	770	807	810	796	752	700	610	497	389	340	345	395			74
72	619	642	680	712	713	719	736	744	779	780	764	717	663	570	453	341	290	295	346			72
70	598	620	658	689	689	695	711	718	751	750	732	682	626	529	409	298	240	245				70
68	577	598	636	666	665	671	686	692	723	720	700	647	589	488	365	245	190					68
66	557	577	614	642	641	647	662	666	695	690	668	612	552	448	321	197	140					66
64	536	555	591	619	616	622	637	639	668	660	636	578	514	407	276	149						64
62	516	534	569	595	592	598	613	613	640	630	604	545	477	367	232							62
60	496	512	547	572	568	574	588	587	612	600	572	508	440	326								60
58	474	490	525	549	544	550	563	561	584	570	540	473	405									58
56	454	469	503	525	520	526	539	535	556	540	508	438										56
54	433	447	480	502	495	501	514	[illegible]	529	510	476											54
52	413	426	458	478	471	477	490	482	501	480												52
50	392	404	436	455	447	453	468	456	[illegible]													50
	8	8-6	9	9-6	10	10-6	11	11-6	12	12-6	13	13-6	14	14-6	15	15-6	16	16-6	17	17-6	18	
*	10.3	10.8	11.1	11.7	12.1	12.1	12.3	13.1	13.9	15	16	17.4	18.6	20.3	22.1	24	25	24.9	24.2	23.3	22.5	

*Weight Deviation Multiplier

Reproduced with permission of Frederick Rand Rogers.

Table A-28
Strength Index Norms for Women (Belt)

AGE

WEIGHT	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	WEIGHT
220													3035	3017	3000	2994	2978	2975	2972	2969	2955	220
218												3017	2998	2981	2964	2958	2942	2939	2936	2933	2919	218
216											3001	2980	2962	2944	2928	2922	2906	2903	2900	2897	2883	216
214										2985	2964	2944	2926	2908	2891	2885	2870	2867	2864	2861	2848	214
212										2948	2927	2907	2889	2871	2855	2849	2834	2831	2828	2825	2812	212
210									2946	2911	2890	2870	2852	2835	2819	2813	2798	2795	2792	2789	2776	210
208								2943	2908	2874	2855	2833	2815	2799	2783	2777	2762	2759	2756	2753	2740	208
206								2905	2871	2837	2816	2796	2779	2762	2747	2741	2726	2723	2720	2717	2704	206
204							2913	2867	2833	2799	2779	2760	2742	2726	2710	2704	2690	2687	2684	2681	2669	204
202							2875	2829	2796	2762	2742	2723	2706	2689	2674	2668	2654	2651	2648	2645	2633	202
200						2873	2836	2791	2758	2725	2705	2686	2669	2653	2638	2632	2618	2615	2612	2609	2597	200
198						2834	2797	2753	2720	2688	2668	2649	2632	2617	2602	2596	2582	2579	2576	2573	2561	198
196					2844	2795	2759	2715	2683	2651	2631	2612	2596	2580	2565	2560	2546	2543	2540	2537	2525	196
194					2804	2756	2720	2677	2645	2613	2594	2576	2559	2544	2529	2523	2510	2507	2504	2501	2490	194
192				2822	2765	2717	2682	2639	2608	2576	2557	2539	2523	2507	2493	2487	2474	2471	2468	2465	2454	192
190				2782	2725	2678	2643	2601	2570	2539	2520	2502	2486	2471	2457	2451	2438	2435	2432	2429	2418	190
188			2808	2742	2685	2639	2604	2563	2532	2502	2483	2465	2449	2435	2421	2415	2402	2399	2396	2393	2382	188
186			2766	2701	2646	2600	2566	2525	2495	2465	2446	2428	2413	2398	2385	2379	2366	2363	2360	2357	2346	186
184		2804	2725	2661	2606	2561	2527	2487	2457	2427	2409	2392	2376	2362	2348	2342	2330	2327	2324	2321	2311	184
182	2880	2761	2683	2620	2567	2522	2489	2449	2420	2390	2372	2355	2340	2325	2312	2306	2294	2291	2288	2285	2275	182
180	2835	2718	2642	2580	2527	2483	2450	2411	2382	2353	2335	2318	2303	2289	2276	2270	2258	2255	2252	2249	2239	180
178	2790	2675	2601	2540	2487	2444	2411	2373	2344	2316	2298	2281	2266	2253	2240	2234	2222	2219	2216	2213	2203	178
176	2745	2632	2559	2499	2448	2405	2373	2335	2307	2279	2261	2244	2230	2216	2204	2198	2186	2183	2180	2177	2167	176
174	2700	2590	2518	2459	2408	2366	2334	2297	2269	2241	2224	2208	2193	2180	2167	2161	2150	2147	2144	2141	2132	174
172	2655	2547	2476	2418	2369	2327	2296	2259	2232	2204	2187	2171	2157	2143	2131	2125	2114	2111	2108	2105	2096	172
170	2610	2504	2435	2378	2329	2288	2257	2221	2194	2167	2150	2134	2120	2107	2095	2089	2078	2075	2072	2069	2060	170
168	2565	2461	2394	2338	2289	2249	2218	2183	2156	2130	2113	2097	2083	2071	2059	2053	2042	2039	2036	2033	2024	168
166	2520	2418	2352	2297	2250	2210	2180	2145	2119	2093	2076	2060	2047	2034	2023	2017	2006	2003	2000	1997	1988	166
164	2475	2376	2311	2257	2210	2171	2141	2107	2081	2055	2039	2024	2010	1998	1986	1980	1970	1967	1964	1961	1953	164
162	2430	2333	2269	2216	2171	2132	2103	2069	2044	2018	2002	1987	1974	1961	1950	1944	1934	1931	1928	1925	1917	162
160	2385	2290	2228	2176	2131	2093	2064	2031	2006	1981	1965	1950	1937	1925	1914	1908	1898	1895	1892	1889	1881	160
158	2340	2247	2187	2136	2091	2054	2025	1993	1968	1944	1928	1913	1900	1889	1878	1872	1862	1859	1856	1853	1845	158
156	2295	2204	2145	2095	2052	2015	1987	1955	1931	1907	1891	1876	1864	1852	1842	1836	1826	1823	1820	1817	1809	156
154	2250	2162	2104	2055	2012	1976	1948	1917	1893	1869	1854	1840	1827	1816	1805	1799	1790	1787	1784	1781	1774	154
152	2205	2119	2062	2014	1973	1937	1910	1879	1856	1832	1817	1803	1791	1779	1769	1763	1754	1751	1748	1745	1738	152
150	2160	2076	2021	1974	1933	1898	1871	1841	1818	1795	1780	1766	1754	1743	1733	1727	1718	1715	1712	1709	1702	150
148	2115	2033	1980	1934	1893	1859	1832	1803	1780	1758	1743	1729	1717	1707	1697	1691	1682	1679	1676	1673	1666	148
146	2070	1990	1938	1893	1854	1820	1794	1765	1743	1721	1706	1692	1681	1670	1661	1655	1646	1643	1640	1637	1630	146
144	2025	1948	1897	1853	1814	1781	1755	1727	1705	1683	1669	1656	1644	1634	1624	1618	1610	1607	1604	1601	1595	144
142	1980	1905	1856	1812	1775	1742	1717	1689	1668	1646	1632	1619	1608	1597	1588	1582	1574	1571	1568	1565	1559	142
140	1935	1862	1814	1772	1735	1703	1678	1651	1630	1609	1595	1582	1571	1561	1552	1546	1538	1535	1532	1529	1523	140
138	1890	1819	1773	1732	1695	1664	1639	1613	1592	1572	1558	1545	1534	1525	1516	1510	1502	1499	1496	1493	1487	138
136	1845	1776	1731	1691	1656	1626	1601	1575	1555	1535	1521	1508	1498	1488	1480	1474	1466	1463	1460	1457	1451	136
134	1800	1734	1690	1651	1616	1586	1562	1537	1517	1497	1484	1472	1461	1452	1443	1437	1430	1427	1424	1421	1416	134
132	1755	1691	1648	1610	1577	1547	1524	1499	1480	1460	1447	1435	1425	1415	1407	1401	1394	1391	1388	1385	1380	132
130	1710	1648	1607	1570	1537	1508	1485	1461	1442	1423	1410	1398	1388	1379	1371	1365	1358	1355	1352	1349	1344	130
128	1665	1605	1566	1530	1497	1469	1446	1423	1404	1386	1373	1361	1351	1343	1335	1329	1322	1319	1316	1313	1308	128
126	1620	1562	1524	1489	1458	1430	1408	1385	1367	1349	1336	1324	1315	1306	1299	1293	1286	1283	1280	1277	1272	126
124	1575	1520	1483	1449	1418	1391	1369	1347	1329	1311	1299	1288	1278	1270	1262	1256	1250	1247	1244	1241	1236	124
122	1530	1477	1441	1408	1379	1352	1331	1309	1292	1274	1262	1251	1242	1233	1226	1220	1214	1211	1208	1205	1201	122
120	1485	1434	1400	1368	1339	1313	1292	1271	1254	1237	1225	1214	1205	1197	1190	1184	1178	1175	1172	1169	1165	120
118	1440	1391	1359	1328	1299	1274	1253	1233	1216	1200	1188	1177	1168	1161	1154	1148	1142	1139	1136	1133	1129	118
116	1395	1348	1317	1287	1260	1235	1215	1195	1179	1163	1151	1140	1132	1124	1118	1112	1106	1103	1100	1097	1093	116
114	1350	1306	1276	1247	1220	1196	1176	1157	1141	1125	1114	1104	1095	1088	1081	1075	1070	1067	1064	1061	1058	114
112	1305	1263	1234	1206	1181	1157	1138	1119	1104	1088	1077	1067	1059	1051	1045	1039	1034	1031	1028	1025	1022	112
110	1260	1220	1193	1166	1141	1118	1099	1081	1066	1051	1040	1030	1022	1015	1009	1003	998	995	992	989	986	110
108	1215	1177	1152	1126	1101	1079	1060	1043	1028	1014	1003	993	985	979	973	967	962	959	956			108
106	1170	1134	1110	1085	1062	1040	1022	1005	991	977	966	956	949	942	937	931	926	923				106
104	1125	1092	1069	1045	1022	1001	983	967	953	939	929	920	912	906	900	894	890					104
102	1080	1049	1027	1004	983	962	945	929	916	902	892	883	876	869	864	858						102
100	1035	1006	986	964	943	923	906	891	878	865	855	846	839	833	828							100
98	990	963	945	924	903	884	867	853	840	828	818	809	802	797								98
96	945	920	903	883	864	845	829	815	803	791	781	772	766									96
94	900	878	862	843	824	806	790	777	766	753	744	736										94
92	855	835	820	802	785	767	752	739	728	716	707											92
90	810	792	779	762	746	728	713	701	690													90
	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	
*	22.5	21.4	20.7	20.2	19.8	19.5	19.3	19	18.8	18.6	18.5	18.4	18.3	18.3	18.1	18.1	18	18	18	18	17.9	

*Weight Deviation Multiplier

Reproduced with permission of Frederick Rand Rogers.

Table A-29
Strength Index Norms for Men (Belt)

WEIGHT \ AGE	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	WEIGHT
230													3845	3796	3755	3723	3690	3664	3642	3621	3601	230
228													3808	3759	3719	3686	3653	3628	3606	3585	3565	228
226													3771	3723	3682	3649	3617	3592	3570	3549	3529	226
224													3734	3686	3646	3618	3581	3555	3534	3518	3498	224
222													3697	3650	3609	3577	3545	3519	3498	3477	3457	222
220												3706	3660	3613	3573	3541	3509	3488	3462	3441	3421	220
218												3669	3623	3576	3537	3504	3473	3447	3426	3405	3385	218
216												3632	3586	3540	3500	3468	3437	3411	3390	3369	3349	216
214											3650	3595	3549	3503	3464	3432	3401	3375	3354	3333	3313	214
212											3613	3557	3512	3467	3427	3396	3365	3339	3318	3297	3277	212
210										3626	3575	3520	3476	3430	3391	3359	3329	3303	3282	3261	3241	210
208										3588	3538	3483	3439	3393	3355	3323	3292	3267	3246	3225	3205	208
206										3550	3500	3446	3402	3357	3318	3287	3256	3231	3210	3189	3169	206
204									3567	3512	3462	3408	3365	3320	3282	3251	3220	3195	3174	3153	3133	204
202								3580	3529	3474	3425	3371	3328	3284	3243	3214	3184	3159	3138	3117	3097	202
200							3594	3541	3490	3436	3387	3334	3291	3247	3209	3178	3148	3123	3102	3081	3061	200
198						3603	3554	3502	3451	3398	3349	3297	3254	3210	3173	3142	3112	3087	3066	3045	3025	198
196					3596	3562	3514	3462	3413	3360	3312	3260	3217	3174	3136	3105	3076	3051	3030	3009	2989	196
194				3564	3554	3521	3474	3423	3374	3322	3274	3222	3180	3137	3100	3069	3040	3015	2994	2973	2953	194
192			3515	3522	3512	3480	3433	3384	3335	3284	3236	3185	3143	3101	3063	3033	3004	2979	2958	2937	2917	192
190			3473	3480	3471	3439	3393	3344	3297	3246	3199	3148	3107	3064	3027	2997	2968	2943	2922	2901	2881	190
188		3412	3431	3438	3429	3398	3353	3305	3258	3208	3161	3111	3070	3027	2991	2960	2931	2907	2886	2865	2845	188
186	3283	3371	3389	3396	3387	3357	3313	3265	3220	3170	3124	3073	3033	2991	2964	2924	2895	2871	2850	2829	2809	186
184	3242	3329	3347	3354	3345	3316	3273	3226	3181	3132	3086	3036	2996	2954	2918	2888	2859	2835	2814	2793	2773	184
182	3201	3287	3306	3312	3304	3275	3233	3187	3142	3094	3048	2999	2959	2918	2881	2852	2823	2799	2778	2757	2737	182
180	3159	3246	3264	3270	3262	3234	3192	3147	3104	3056	3011	2962	2922	2881	2845	2815	2787	2763	2742	2721	2701	180
178	3118	3204	3222	3228	3220	3193	3152	3108	3065	3018	2973	2924	2885	2844	2809	2779	2751	2727	2706	2685	2665	178
176	3077	3162	3180	3186	3179	3152	3112	3069	3026	2980	2935	2887	2848	2808	2772	2743	2715	2691	2670	2649	2629	176
174	3035	3121	3138	3144	3137	3112	3072	3029	2988	2942	2898	2850	2811	2771	2736	2707	2679	2654	2634	2613	2593	174
172	2994	3079	3096	3102	3095	3071	3032	2990	2949	2904	2860	2813	2774	2735	2699	2670	2643	2618	2598	2577	2557	172
170	2953	3037	3055	3060	3054	3030	2992	2951	2910	2866	2822	2775	2738	2698	2663	2634	2607	2582	2562	2541	2521	170
168	2911	2996	3013	3018	3012	2989	2951	2911	2872	2828	2785	2738	2701	2661	2627	2598	2570	2546	2526	2506	2485	168
166	2870	2954	2971	2976	2970	2948	2911	2872	2833	2790	2747	2701	2664	2625	2590	2562	2534	2510	2490	2469	2449	166
164	2829	2912	2929	2934	2928	2907	2871	2833	2794	2752	2709	2664	2627	2588	2554	2525	2498	2474	2454	2433	2413	164
162	2787	2871	2887	2892	2887	2866	2831	2793	2756	2714	2672	2626	2590	2552	2517	2489	2462	2438	2418	2397	2377	162
160	2746	2829	2845	2850	2845	2825	2791	2754	2717	2676	2634	2589	2553	2515	2481	2453	2426	2402	2382	2361	2341	160
158	2704	2788	2804	2808	2803	2784	2751	2714	2679	2638	2597	2552	2516	2478	2445	2417	2390	2366	2346	2325	2305	158
156	2663	2746	2762	2766	2762	2743	2710	2675	2640	2600	2559	2515	2479	2442	2408	2380	2354	2330	2310	2289	2269	156
154	2622	2704	2720	2724	2720	2702	2670	2636	2601	2562	2521	2477	2442	2405	2372	2344	2318	2294	2274	2253	2233	154
152	2580	2663	2678	2682	2678	2661	2630	2596	2563	2524	2484	2440	2405	2369	2335	2308	2282	2258	2238	2217	2197	152
150	2539	2621	2636	2640	2637	2620	2590	2557	2524	2486	2446	2403	2369	2332	2299	2272	2246	2222	2202	2181	2161	150
148	2498	2579	2594	2598	2595	2579	2550	2518	2485	2448	2408	2366	2332	2295	2263	2235	2209	2186	2165	2145	2125	148
146	2456	2538	2553	2556	2553	2538	2510	2478	2447	2410	2371	2329	2295	2259	2226	2199	2173	2150	2129	2109	2089	146
144	2415	2496	2511	2514	2511	2497	2470	2439	2408	2372	2333	2291	2258	2222	2190	2163	2137	2114	2093	2073	2053	144
142	2374	2454	2469	2472	2470	2456	2429	2400	2369	2334	2295	2254	2221	2186	2153	2126	2101	2078	2057	2037	2017	142
140	2332	2413	2427	2430	2428	2415	2389	2360	2331	2296	2258	2217	2184	2149	2117	2090	2065	2042	2021	2001	1981	140
138	2291	2371	2385	2388	2386	2374	2349	2321	2292	2258	2220	2180	2147	2112	2081	2054	2029	2006	1985	1966	1945	138
136	2249	2330	2343	2346	2345	2333	2309	2281	2254	2220	2183	2142	2110	2076	2044	2018	1993	1970	1949	1929	1909	136
134	2208	2288	2301	2304	2303	2292	2269	2242	2215	2182	2145	2105	2073	2039	2008	1981	1957	1934	1913	1893	1873	134
132	2167	2246	2260	2262	2261	2251	2229	2203	2176	2144	2107	2068	2036	2003	1971	1945	1921	1898	1877	1857	1837	132
130	2125	2205	2218	2220	2220	2210	2188	2163	2138	2106	2070	2031	2000	1966	1935	1909	1885	1862	1841	1821	1801	130
128	2084	2163	2176	2178	2178	2169	2148	2124	2099	2068	2032	1993	1963	1929	1899	1873	1848	1826	1805	1785	1765	128
126	2043	2121	2134	2136	2136	2128	2108	2085	2060	2030	1994	1956	1926	1893	1862	1836	1812	1790	1769	1749	1729	126
124	2001	2080	2092	2094	2094	2088	2068	2045	2022	1992	1957	1919	1889	1856	1826	1800	1776	1753	1733	1713	1693	124
122	1960	2038	2050	2052	2053	2047	2028	2006	1983	1954	1919	1882	1852	1820	1789	1764	1740	1717	1697	1677	1657	122
120	1919	1996	2009	2010	2011	2006	1988	1967	1944	1916	1881	1844	1815	1783	1753	1728	1704	1681	1661	1641	1621	120
118	1877	1955	1967	1968	1969	1965	1947	1927	1906	1878	1844	1807	1778	1746	1717	1691	1668	1645	1625	1605	1585	118
116	1836	1913	1925	1926	1928	1924	1907	1888	1867	1840	1806	1770	1741	1710	1680	1655	1632	1609	1589	1569		116
114	1795	1871	1883	1884	1886	1883	1867	1849	1828	1802	1768	1733	1704	1673	1644	1619	1596	1573	1553			114
112	1753	1830	1841	1842	1844	1842	1827	1809	1790	1764	1731	1695	1667	1637	1607	1583	1560	1537				112
110	1712	1788	1799	1800	1803	1801	1787	1770	1751	1726	1695	1658	1631	1600	1571	1546	1524					110
108	1670	1747	1758	1758	1761	1760	1747	1730	1713	1688	1656	1621	1594	1563	1535	1510						108
106	1629	1705	1716	1716	1719	1719	1706	1691	1674	1650	1618	1584	1557	1527	1498							106
104	1588	1663	1674	1674	1677	1678	1666	1652	1635	1612	1580	1546	1520	1490								104
102	1546	1622	1632	1632	1636	1637	1626	1612	1597	1574	1543	1509	1483	1454								102
100	1505	1580	1590	1590	1594	1596	1586	1573	1558	1536	1505	1472	1446									100
	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	
*	20.68	20.82	20.92	21	20.85	20.46	20.08	19.68	19.32	19	18.82	18.62	18.45	18.30	18.20	18.13	18.06	18.02	18.01	18	18	

*Weight Deviation Multiplier

1. Norms for individuals whose weights are above limits for which norms are included are calculated by adding to the norm for any che weight the pound difference between that weight and the individual's weight times the Weight Deviation Multiplier.
2. Norms for men over 38 years of age may be calculated roughly by subtracting twenty points from the 38-year norm for each year over 3

Appendix B

CHECK LIST ON SCHOOL ATHLETICS*

Purposes of School Athletics

1. Does your school have clearly defined goals for its athletic program?
   Y....... N....... U....... A....... D....... U.......
2. Do other teachers, as well as coaches and teachers of physical education, have a part in formulating the purposes of athletics in your school?
   Y....... N....... U....... A....... D....... U.......
3. Do lay citizens have opportunity to express themselves with respect to the purposes of athletics in your school?
   Y....... N....... U....... A....... D....... U.......
4. Are athletics recognized by your school as an integral part of complete education?
   Y....... N....... U....... A....... D....... U.......
5. Are athletic activities in harmony with the objectives of the total educational program?
   Y....... N....... U....... A....... D....... U.......
6. Does your school's athletic program encourage participation in satisfying play by *all* pupils?
   Y....... N....... U....... A....... D....... U.......
7. Does your school in its athletic program seek to contribute to the development of wholesome personalities?
   Y....... N....... U....... A....... D....... U.......
8. Does your school conduct athletics in ways intended to help participants develop health and physical fitness?
   Y....... N....... U....... A....... D....... U.......
9. Does your school seek to conduct athletics in such a way that participants develop enduring play habits, skills, and attitudes?
   Y....... N....... U....... A....... D....... U.......
10. Does your school encourage athletic activities which aid development of desirable social growth and adjustment?
    Y....... N....... U....... A....... D....... U.......
11. Does your school conduct athletics in such a way as to avoid excessive emotional strains and tensions on the part of both players and spectators?
    Y....... N....... U....... A....... D....... U.......

* Permission granted to reproduce this check list from *School Athletics: Problems and Policies,* Educational Policies Commission, Washington, 1953.

12. Does your school take steps to prevent athletic practices which might be detrimental to the welfare of pupils as individuals?

Y....... N....... U....... A....... D....... U.......

Health and Welfare of Athletic Participants

13. In determining policies and procedures for athletics in your school is the health and welfare of participants considered paramount?

Y....... N....... U....... A....... D....... U.......

14. Does the school provide adequate protective equipment and other health safeguards for all participants in athletic contests?

Y....... N....... U....... A....... D....... U.......

15. Are all games and practice sessions conducted in facilities that are hygienic, clean, and safe?

Y....... N....... U....... A....... D....... U.......

16. Is adequate training and conditioning required for all types of athletic competition?

Y....... N....... U....... A....... D....... U.......

17. Has the approval of a physician been secured for the practices of your school's interscholastic athletic program which involve conditioning, training, and health?

Y....... N....... U....... A....... D....... U.......

18. Is a thorough health examination required of all participants in both intramural and interscholastic sports before they take part in vigorous athletic competition?

Y....... N....... U....... A....... D....... U.......

19. Is a postseason health examination required of athletes?

Y....... N....... U....... A....... D....... U.......

20. Is emergency medical service available during all practice periods, intramural games, and interscholastic contests held under school auspices?

Y....... N....... U....... A....... D....... U.......

21. Are athletes who have been injured or ill readmitted to participation only with the written approval of a physician?

Y....... N....... U....... A....... D....... U.......

22. Does your school have a written and well-publicized policy regarding the legal and financial responsibilities for injuries incurred in athletics?

Y....... N....... U....... A....... D....... U.......

23. Even when not legally responsible, does your school have a plan for making financial provisions for the care of injuries incurred in school athletics?

Y....... N....... U....... A....... D....... U.......

24. Does the school seek to prevent injury to the personality development of star athletes from overattention and ego-inflation?

Y....... N....... U....... A....... D....... U.......

Organization and Administration of School Athletics

25. Are all athletic activities in your school recognized as the responsibility of the school and under its control?

Y....... N....... U....... A....... D....... U.......

26. Are all athletics in your school administered as part of the school's total program of physical education?

Y....... N....... U....... A....... D....... U.......

27. Is your school (if a high school) a member of your state high school athletic association or similar organization?

Y....... N....... U....... A....... D....... U.......

28. Does your school accept the aid of your state's department of education (or public instruction) in establishing and maintaining high standards in the conduct of school athletics?

Y....... N....... U....... A....... D....... U...

Facilities for School Athletics

29. Does your board of education provide adequate facilities in athletics for *all* students?

Y....... N....... U....... A....... D....... U.......

30. Are physical education facilities in your school available to all phases of the program, including required activity classes and intramurals?

Y....... N....... U....... A....... D....... U.......

31. Do girls share equally with boys in the use of your school's athletic facilities?

Y....... N....... U....... A....... D....... U.......

32. Does your school provide a standard field, court, or play space for each team game and individual sport most popular in your section of the country?

Y....... N....... U....... A....... D....... U.......

Personnel for the Athletic Program

33. Are all who coach athletic teams in your school competently trained and certified as teachers?

Y....... N....... U....... A....... D....... U.......

34. Do the athletic coaches have professional training in physical education equivalent to a minor or more?

Y....... N....... U....... A....... D....... U.......

35. Do the athletic coaches consistently set good examples in the matter of sportsmanship and personal conduct?

Y....... N....... U....... A....... D....... U.......

36. Do athletic coaches use their influence with students to help them with personal problems?

Y....... N....... U....... A....... D....... U.......

37. Are those members of the school staff whose chief work is coaching athletics generally regarded by other faculty members as fellow teachers of comparable professional status?

Y....... N....... U....... A....... D....... U.......

38. Does the school provide sufficient personnel for the proper instruction and supervision of all participants in the required activity classes, in co-recreation, in intramural sports, and in interscholastic athletics?

Y....... N....... U....... A....... D....... U.......

39. Do school authorities seek to maintain at all times a balance in the amount of staff time and instruction given to all phases of physical education, including

required activity classes, co-recreation, intramural sports, and interscholastic athletics?

Y....... N....... U....... A....... D....... U.......

Intramural Programs

40. Does every student in your school system have opportunity for participating in a variety of intramural sports?

Y....... N....... U....... A....... D....... U.......

41. Is the intramural sports program conducted as an integral part of the total program of physical education and not as a "feeder" system for interscholastic athletics?

Y....... N....... U....... A....... D....... U.......

42. Does the intramural sports program serve as a laboratory where students can test the things they are taught in physical education classes?

Y....... N....... U....... A....... D....... U.......

43. Do most of the students in your school find the intramural program sufficiently interesting, diverse, and convenient that they voluntarily participate in it?

Y....... N....... U....... A....... D....... U.......

44. Does the school provide opportunities for co-recreation (that is, for boys and girls to play together) through intramural sports?

Y....... N....... U....... A....... D....... U.......

45. Does the intramural athletic program have good equipment rather than handed-down equipment, worn-out balls, unmarked fields, and poorly organized game situations?

Y....... N....... U....... A....... D....... U.......

Elementary School Policies and Programs

46. Does your elementary school program recognize that children need to engage in vigorous muscular activities as a part of the growing process?

Y....... N....... U....... A....... D....... U.......

47. Do the practices of your school reflect the policy that interscholastic athletics are not desirable in elementary schools?

Y....... N....... U....... A....... D....... U.......

48. Are sports adapted to the needs of children of elementary school age by modification of such things as rules, size of courts and fields, and the size or weight of equipment?

Y....... N....... U....... A....... D....... U.......

49. Are athletic activities in your elementary school planned cooperatively by teachers, pupils, and parents?

Y....... N....... U....... A....... D....... U.......

50. Are children of elementary school age given opportunities to develop responsibility through the athletic program by participating in the planning, by organizing groups, by holding office, and by helping with equipment?

Y....... N....... U....... A....... D....... U.......

51. Do the school authorities discourage and seek to prevent outside interests from forcing undesirable patterns of athletic competition upon children?

Y....... N....... U....... A....... D....... U.......

Junior High School Policies and Programs

52. Is the athletic program for junior high school pupils suited to the needs of children who are undergoing rapid changes in physical growth?

    Y....... N....... U....... A....... D....... U.......

53. Is the athletic program of your junior high school planned cooperatively by teachers, pupils, and parents?

    Y....... N....... U....... A....... D....... U.......

54. Does your junior high school provide a broad athletic program for every boy and girl in the school?

    Y....... N....... U....... A....... D....... U.......

55. Does the athletic program in junior high school consist primarily of sports organized and conducted on an intramural basis?

    Y....... N....... U....... A....... D....... U.......

56. Are junior high school pupils given opportunities to develop responsibility through the athletic program by participating in the planning, by organizing groups, by holding office, and by helping with equipment?

    Y....... N....... U....... A....... D....... U.......

57. Is tackle football prohibited as an athletic activity in your junior high school?

    Y....... N....... U....... A....... D....... U.......

58. Does your junior school refrain from participation in varsity-type interscholastics?

    Y....... N....... U....... A....... D....... U.......

59. Are the leaders of athletic activities in your junior high school competently trained teachers?

    Y....... N....... U....... A....... D....... U.......

Interscholastic Athletics for Boys in Senior High School

60. Are interscholastic athletics conducted primarily to serve the needs of students with superior athletic skills?

    Y....... N....... U....... A....... D....... U.......

61. Are games and practice periods for interscholastic athletics worked into the school schedule with a minimum of interference with the academic program?

    Y....... N....... U....... A....... D....... U.......

62. Are interscholastic games played only with schools that maintain acceptable principles and policies in their conduct of interscholastic athletics?

    Y....... N....... U....... A....... D....... U.......

63. Are interscholastic games played only on school or public property?

    Y....... N....... U....... A....... D....... U.......

64. Does the school observe the rules of its state high school athletic association?

    Y....... N....... U....... A....... D....... U.......

65. Does the school refuse to participate in all postseason tournaments and post-season championship games?

    Y....... N....... U....... A....... D....... U.......

66. Is the board of education adequately informed regarding the interscholastic athletic program?

    Y....... N....... U....... A....... D....... U.......

67. Is the board of education adequately informed regarding the rules and regulations of the state high school athletic association?

    Y....... N....... U....... A....... D....... U.....

68. Are athletes engaged in interscholastic sports held to the same standards of scholarship as other students?

Y....... N....... U....... A....... D....... U.......

69. Are boys who participate in interscholastic athletics required to attend regular classes in physical education (except during the actual period of their interscholastic participation)?

Y....... N....... U....... A....... D....... U.......

70. Does the school try to prevent solicitation of its athletes by colleges and universities through tryouts and competitive bidding?

Y....... N....... U....... A....... D....... U.......

71. Does the school make an effort to develop high standards of good sportsmanship on the part of all students?

Y....... N....... U....... A....... D....... U.......

Athletics for Girls

72. Does the school athletic program for girls provide opportunities for all girls to participate according to their needs, abilities, and interests?

Y....... N....... U....... A....... D....... U.......

73. Are facilities for girls' athletics provided in accordance with the requirements of the girls' program and not on the basis of causing minimum inconvenience to the boys' program?

Y....... N....... U....... A....... D....... U.......

74. Is the school athletic program for girls under the direction of a competent woman leader who is professionally trained in health and physical education?

Y....... N....... U....... A....... D....... U.......

75. Does the school include in its girls' athletic program such dual, individual, and recreational sports and games as archery, badminton, bowling, croquet, horseshoes, golf, riding, shuffleboard, skiing, swimming, table tennis, and tennis?

Y....... N....... U....... A....... D....... U.......

76. Does the school include in its girls' athletic program such team sports as basketball, field hockey, softball, soccer, speedball, and volleyball?

Y....... N....... U....... A....... D....... U.......

77. Does every girl in the school have an opportunity to compete in team games?

Y....... N....... U....... A....... D....... U.......

78. Is participation of girls in athletics based on an appraisal of the health status of each participant which takes into account quality and extent of participation, type of activity, individual differences, and general organic condition?

Y....... N....... U....... A....... D....... U.......

79. Are girls prevented from participation, under school auspices, in sports which involve rough and tumble body contact?

Y....... N....... U....... A....... D....... U.......

80. Are girls provided opportunities for athletic competition with girls of other schools through such means as play days and sports days?

Y....... N....... U....... A....... D....... U.......

81. Are all school athletic contests for girls conducted in accordance with girls' rules?

Y....... N....... U....... A....... D....... U.......

82. Is the school's athletic program for girls conducted in conformity with the policies and recommendations contained in *Standards in Sports for Girls and Women?*

Y....... N....... U....... A....... D....... U.......

Financing Athletic Programs

83. Does the board of education control the financing of the athletic program?
    Y....... N....... U....... A....... D....... U.......
84. Is the welfare of participants considered more important than financial gain in determining the athletic policies and practices of your school?
    Y....... N....... U....... A....... D....... U.......
85. Is the size of the budget for athletics in sound proportion to the size of the budget for the rest of the school program?
    Y....... N....... U....... A....... D....... U.......
86. Is balance maintained in the financial support of all phases of the physical education program, including required activity classes, co-recreation, intramural sports, and interscholastic athletics?
    Y....... N....... U....... A....... D....... U.......
87. Are the salaries of other teachers equitable in comparison with the salaries of coaches?
    Y....... N....... U....... A....... D....... U.......
88. Are all athletic moneys, including gate receipts, considered as school funds with records accurately kept and audited?
    Y....... N....... U....... A....... D....... U.......
89. Is the intramural sports program of your school financed entirely by appropriations from tax funds?
    Y....... N....... U....... A....... D....... U.......
90. Is your school district moving toward complete financing of the athletic program from tax funds?
    Y....... N....... U....... A....... D....... U.......

Community Relations

91. Is the school actively concerned with providing adequate community athletic facilities for children and youth?
    Y....... N....... U....... A....... D....... U.......
92. Does the school staff study your community to determine how to use, to improve, and to increase available space for wholesome play for the students enrolled in your school?
    Y....... N....... U....... A....... D....... U.......
93. Is the school actively interested in providing community athletic programs for children and youth during vacation periods?
    Y....... N....... U....... A....... D....... U.......
94. Do community organizations look to the school for expert counsel and advice regarding athletic programs?
    Y....... N....... U....... A....... D....... U.......
95. Does the school assume responsibility for informing the community regarding the standards of good sportsmanship that should be observed at all athletic contests?
    Y....... N....... U....... A....... D....... U.......
96. Is the conduct of spectators at interscholastic games such as to reflect favorably on the school?
    Y....... N....... U....... A....... D....... U.......

97. Does the board of education enjoy as much freedom from outside pressures in the selection of a coach as it does in the selection of other teachers?

Y....... N....... U....... A....... D....... U.......

98. Does the community support the coach of a boys' varsity team that has a losing season?

Y....... N....... U....... A....... D....... U.......

99. Do local newspapers, radio stations, and television stations support clean athletics and have a high regard for standards of good sportsmanship?

Y....... N....... U....... A....... D....... U.......

100. Are the schools comparatively free from undesirable activities on the part of outside organizations primarily concerned with winning teams in interscholastic games?

Y....... N....... U....... A....... D....... U.......

HEWITT COMPREHENSIVE TENNIS KNOWLEDGE TEST (FORM A)

Directions About half of the following statements are true and about one half are false. Place a plus (+) sign before statement if true, and if the statement is partially or wholly false put (0) before the statement. Mark statements in order. Do not guess, for guessing reduces your score. Answer all questions as for a right-handed player.

(1)______A chop stroke produces backward spin on the ball.

(2)______One should attempt to hit a ball on the forehand drive at the level of the shoulder or eyes.

(3)______Hitting the ball in the air before it touches the ground is called a half-volley.

(4)______If time permits, a pause at the end of the backward swing on all drives is advisable.

(5)______Putting top spin on the ball makes it stay in the air longer.

(6)______The end line of the tennis court is called the back line.

(7)______It is good tennis strategy to play through strength to weakness.

(8)______To step on the line while serving is called a line fault.

(9)______A "smash" is an attempt to "kill" the ball on a forehand or backhand drive from a ground bounce.

(10)______To lob the ball in the sun is to take an unfair advantage of the opponent.

(11)______In the backhand drive the left shoulder points toward and is at right angles to the net.

(12)______Racket stringing can be too tight for good play.

(13)______There are only two phases of the forehand drive—the backward swing and hitting the ball.

(14)______In the reverse American twist serve, the ball bounces toward the opponent's right.

(15)______"Wrightman Cup" play is competition for women only in countries of England and the United States.

Scoring R − W* Score______________

Directions Place in the parentheses the number of the word or phrase that best completes the sentence.

(16) ( ) It is good strategy in tennis to: (1) force opponent out of position, (2) drive all balls to opponent's backhand side, (3) try to rattle opponent during play, (4) play balls to opponent's baseline, (5) chop all balls.

(17) ( ) In the chop stroke where the ball is hit hip high, the racket completes the follow through· (1) above head, (2) above the eyes, (3) below the eyes, (4) below the shoulders, (5) below the hips.

(18) ( ) To be at the net position first is: (1) spectacular, (2) best offensive position, (3) worst offensive position, (4) best lobbing position, (5) best forehand drive position.

(19) ( ) In the forehand drive, one should attempt to contact the ball: (1) in front of fore foot, (2) even with fore foot, (3) center of body, (4) even with rear foot, (5) behind rear foot.

(20) ( ) The overhead volley has somewhat the same motions as the smash with: (1) more speed, (2) more cut, (3) less speed, (4) less cut, (5) more wrist action.

(21) ( ) If the opponent hits to your weakness you should: (1) run around the ball

and hit it on your strong side, (2) lob, (3) play off to one side so that the ball comes to your strong side, (4) make attempt to play the ball on the side it hits.

(22) ( ) The best position to cover most lobs on the court is: (1) 2 feet from net, (2) midway between the baseline and net, (3) back of baseline, (4) alleys, (5) midway between baseline and service line.

(23) ( ) Smashing is more liable to be successful at the: (1) spot one foot from net, (2) back of the service line, (3) baseline, (4) back of the service line and in alleys, (5) spot between the net and the service line.

Scoring $\frac{R - W}{N - 1}$ Score______

Directions Enumerate the names for following lines and spaces of a tennis court.

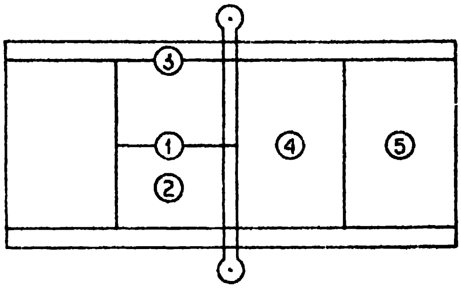
(24)______________________(1)

(25)______________________(2)

(26)______________________(3)

(27)______________________(4)

(28)______________________(5)



Scoring Number right Score______________

Directions Answer the questions below by placing either "Yes" or "No" on the line before each question.

(29)______Player A served underhand. Umpire said nothing. Was this correct?

(30)______Player B was in such a position that he played the ball with both hands on the racket. Player A did not play the return, but claimed the point. Umpire gave Player A the point because B used two hands. Was this correct?

(31)______Player B throws up the second ball poorly on the serve and instead of hitting the ball he catches it instead. A claims the point because B caught the ball. Was this right?

(32)______Is the score called "game" when one side wins the additional point after deuce?

(33)______Player F cuts the ball over the net. It has so much backspin on it that the ball bounces back into F's own court. D comes up fast to the net, but does not touch the net. He reaches over and touches the ball before it bounces a second time. F then knocks the ball out of bounds, D was given the point. Was this correct?

(34)______In tournament play for women does winning three out of five sets determine the match?

(35)______Player C runs close to the net and in returning the ball finds that he must jump the net to keep from touching the net. Player B, the opponent, becomes confused and knocks the ball out of the court. Umpire gave point to C. Was this correct?

(36)______Is it permissible for your partner in a doubles game to stand in the center of the court obstructing the view of the receiver?

(37)______If the first served ball is still in the service court, and you fail to remove it, and the second served ball hits it, is the point scored for opponent?

(38)______In tournament play are courts changed at the end of every even game played?

(39)______A makes good return at the net. In so doing his racket slips from his hand and into B's court. B because of this becomes confused and misses the ball. Does B get the point?

(40)______Is the best net position midway between the service line and the baseline?

(41)______Player C standing close to the net. Player D hits the ball close to C. Before the ball passes over the net, C reaches over the net and touches the ball with his racket. D could not return it. C, however, did not touch the net in reaching over. Was the unpire correct in giving C the point?

(42)______Do cement courts make a faster game of tennis than grass courts?

(43)______Player A, while serving, steps on the line simultaneously as the racket touches the ball. Is this a violation of the rules?

Scoring Number right − Number wrong Score__________

Directions Match the following numbers in Column A with the statement in Column B to which they are related. Have no more than one number before any item in Column B.

Column A

(1) Out	(15) Allison
(2) Kill	(16) Let
(3) Fifteen All	(17) Chop
(4) Game	(18) Continue Play
(5) Advantage Receiver	(19) Reverse Ball
(6) Cahill	(20) Fault
(7) Tilden	(21) Footfault
(8) Smash	(22) Lenglen
(9) Jacobs	(23) Service Line
(10) Advantage Server	(24) Perry
(11) Moody	(25) Lob
(12) Love All	(26) Permissible
(13) Vines	(27) Lott
(14) Thirty All	(28) Service
	(29) Deuce

Column B

(44)______Knocking the ball over the head of the opponent.

(45)______Ball hits net on the serve and falls into the correct service court.

(46)______Both sides winning one point each.

(47)______Winning point after deuce while serving.

(48)______American man tennis player who was famous for his cannon ball service and is now a leading professional.

(49)______Ball spinning backward to its flight.

(50)______Any served ball hitting outside legal boundary.

Scoring Number right Score______

Score Summary

(1) True-False __________

(2) Multiple Choice __________

(3) Diagrammatic __________

(4) Yes and No __________

(5) Matching __________

Final Score __________

*Symbols: R = right; W = wrong; N = number of questions.

SOURCE: Jack E. Hewitt, "Comprehensive Tennis Knowledge Test," *The Research Quarterly,* 8:74–80, October, 1937; Used by permission.

Glossary

Achievement tests These are measures of performance in tests of skill or knowledge; they are distinguished from tests of fitness or personality. Test results are often combined into achievement scales.

Agility The ability to move with both speed and accuracy.

Audiometer An instrument for measuring the power of hearing or the audibility or intensity of sounds.

Bar graph A bar or column used to display comparison between variables where differences exist in one dimension. It is useful for quick identification of results.

Biserial correlation A correlation method used when one of the variables is expressed in two categories.

Central tendency The score in a group of scores that is representative of the total group. In education this score is usually the mode, median, or mean.

Classification index A measure used in the homogeneous grouping of students. Some of factors inherent in this index may involve height, weight, age, motor ability, and specific skills. Several workers have their own classification index.

Coefficient of correlation An index of the degree of relationship of two sets of scores. (See correlation.)

Constitutional psychology The study that treats human behavior as a partial manifestation of constitutional endowment. In short, man behaves as he does not only because of environmental training but also because of what he is made of.

Correlation The relation of two or more sets of paired data. Perfect positive correlation is expressed as 1.0; perfect negative correlation as −1.0; and absence of correlation as 0.0.

Criterion The standard by which validity may be determined. A test or a single factor may be measured or judged against a criterion.

Decibel The unit for measuring the relative loudness of sounds. It is the smallest degree of difference of loudness ordinarily detectable by the human ear, the range of which is about 130 decibels.

Derived score A score obtained by converting a raw score on a test to another figure based on such items as age, group, or percentile.

Difficulty rating A rating given to test questions to indicate the degree to which one question is more difficult than another.

Discriminative index (index of discrimination) A measure of the effectiveness with which a test item discriminates between students of different levels of information or ability.

Education The process of changing behavior toward certain preconceived aims and objectives.

Empirical method A method that relies on practical experience and observations alone. Sometimes tests are established without due consideration of science and theory, which makes them of questionable validity. When the empirical method is employed, one relies almost entirely on what one observes at the moment and it is often assumed that such observations are typical.

Endurance The capacity to sustain an activity over a period of time. This may involve repetition of a movement, such as chinning the bar, or sustained activity, such as holding a heavy weight.

Evaluation Appraisal of the results of measurement in the light of preconceived aims and objectives. It concerns the proper value to be placed on test results.

Frequency distribution A table of scores set up by intervals, showing the distribution of scores in each interval. It is used when further statistics are to be employed.

Guessed mean The mid-point of the class or step interval that is assumed to be the mean when the mean is determined by the short-cut method.

Health education The sum of experiences that favorably influence habits, attitudes, and knowledge relating to individual, community, and racial health.

Health service Those procedures designed to determine the health status of the child, to enlist his cooperation in health protection and maintenance, to inform parents of the defects that may be present, to prevent disease, and to correct remediable defects.

Histogram A graph on coordinate paper where scores or cases in a given interval are represented by a rectangle, the height of which is equal to the number of cases within the interval.

Line graph A line on coordinate paper designed to show the rise and fall of a variable over a period of time.

Mean The point in a distribution of scores usually referred to as the average. It is the point about which the sum of the plus deviations is equal to the sum of the minus deviations.

Measurement The process of finding answers to the questions of how much, how many, and how often. It is concerned with quantities and qualities in evidence.

Median The point in a distribution above which and below which lie 50 per cent of the cases.

Mode The score that appears most frequently and is typical of the other scores in a distribution.

Multiple correlation A correlation method used to relate more than two variables to each other. For example, height, weight, and tennis ability might all be related to each other through correlation.

Norm A standard point of reference that can provide a basis for judgment. It is used by the educator and research worker because the process of measuring involves comparisons.

Normal curve The bell-shaped distribution curve that depicts a massing of scores or data at the center with gradually diminishing numbers toward both extremes.

Objectivity The degree of uniformity with which various persons score the same tests. This is often expressed by a correlation coefficient (r).

Organic efficiency The relative state of the cardiorespiratory mechanism; more generally it can refer to the total physiological activity.

Partial correlation A correlation method used to find the relationship between two variables when other variables, which may affect the true correlation by raising or lowering it, have been ruled out or held constant. For example, weight might be correlated with tennis ability without considering height. Height, therefore, would be held constant, and the partial correlation method would be employed.

Pearson product-moment coefficient of correlation The r method of correlation that seeks a straight-line relationship from the scatter or spread of two or more groups of scores.

Percentile One of the points that divide a distribution into 100 equal parts. Thus a percentile of 62 indicates the point below which lie 62 per cent of the cases.

Percentile graph The curve (ogive curve) illustrating the number of cases or per cent of the population in a distribution of scores lying above or below a given mark.

Physical education The process of changing behavior toward certain preconceived goals primarily through large muscle activity.

Physical fitness This is essentially physiological capacity for activity. It is bodily strength, vigor, and endurance—the wherewithal for all movement related to physical, mental, and emotional status.

Physical fitness index A score derived from dividing the gross score on several tests of strength and muscular endurance by a norm for the subject's height and weight. A score of 100 is the median nationally, but pupils in any school may range from PFI 40 or less to PFI 160 or more.

Predictive index This index ($1 - \sqrt{1 - r^2}$) is used to transform the coefficient of correlation (r) to a better than chance figure. It is given in per cent.

Profile graph A distinctive outline (contour) of a group of scores or qualities usually indicating the student's level of achievement

Quartile deviation (Q) This is the semi-interquartile range that is half the range of the middle half of the scores.

Range The interval between the lowest and highest scores in a distribution. If the lowest score is 8 and the highest score is 94, the range is 86.

Rank-order correlation (Spearman) A correlation method designated as rho (ρ) and used with a small number of cases. Scores are ranked in two series and differences are noted.

Rating scale A measure used to ascertain the relative standing of a person or object. It may be used for classification purposes or to determine status.

Raw score The basic test score. It is usually changed to a derived score for purposes of comparing test data.

Reliability The degree to which a test is accurate, that is, will repeat itself. All measures vary in reliability owing to the nature of the sample, the measuring process, and the subject being tested.

Sampling Selecting subjects or scores at random from a large group or population so that they have the characteristics of the whole population.

Scattergram A table showing the scatter of two distributions at the same time and the relationship that exists between the two.

School health education That part of health education that takes place in school or through efforts organized and conducted by school personnel.

Sigma A Greek letter (σ) used to represent the standard deviation.

Skewed distribution An assymetrical distribution of scores.

Somatotype A physical type of man. This is part of the Sheldon method of classification based on the structure, functions, and development of the human constitution. (See constitutional psychology.)

Spearman-Brown prophecy formula A formula used with two tests that are similar. The reliability of a given test (which is longer or shorter than a known test) is checked against the known test reliability.

Standard deviation A reliable measure of variability obtained by deriving the square root of the mean of the squares of all the deviations from the arithmetic mean.

Strength index A score derived from combining the scores on tests of strength, muscular endurance, and vital capacity.

T score A score based on a 100-point scale and figured on the standard deviation values of a distribution. The mean T score is 50, and each 10 points above and below represent one standard deviation.

Variability The scatter or spread of scores within a distribution of scores.

Validity The degree to which a test measures what it is supposed to measure.

Z score A score derived by subtracting the raw score from the mean and dividing by the standard deviation.

Index